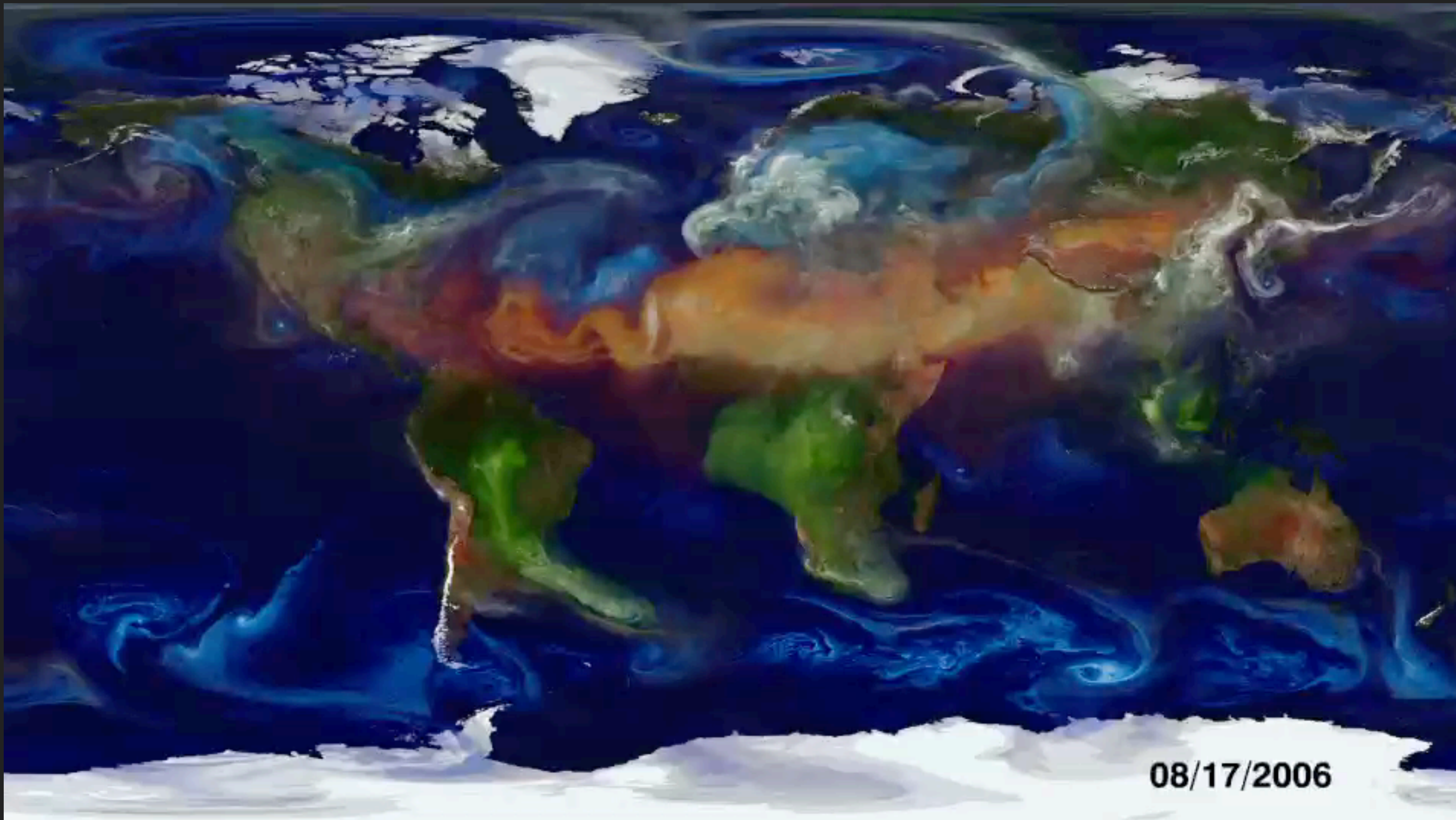


CESM2 DEVELOPMENT AS VIEWED THROUGH THE LENS OF THE NCAR CLIMATE MODEL ANALYSIS TOOL (CMAT)

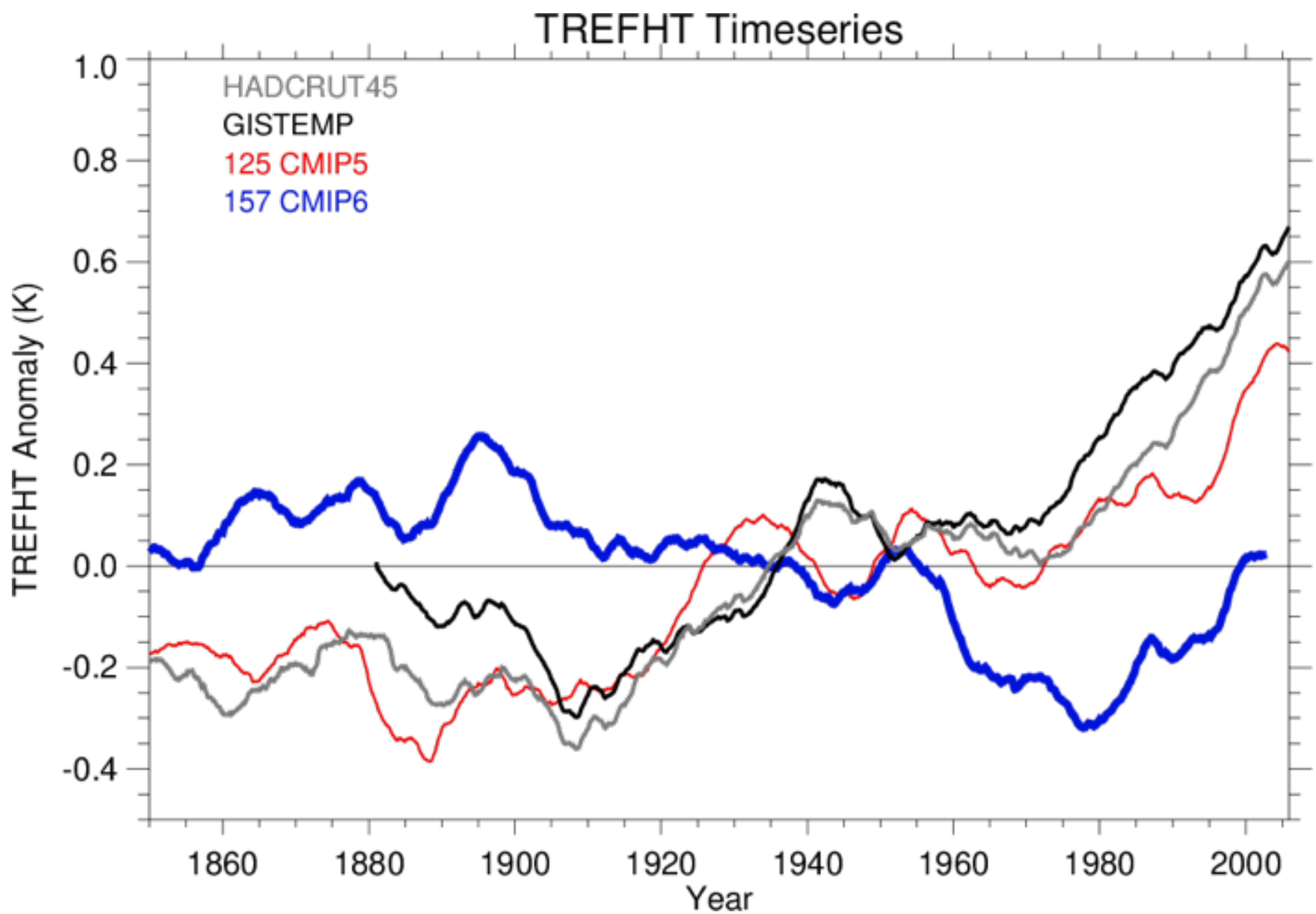
J. FASULLO

2018 Earth Radiation Budget Workshop, NCAR, Boulder CO

OUTLINE

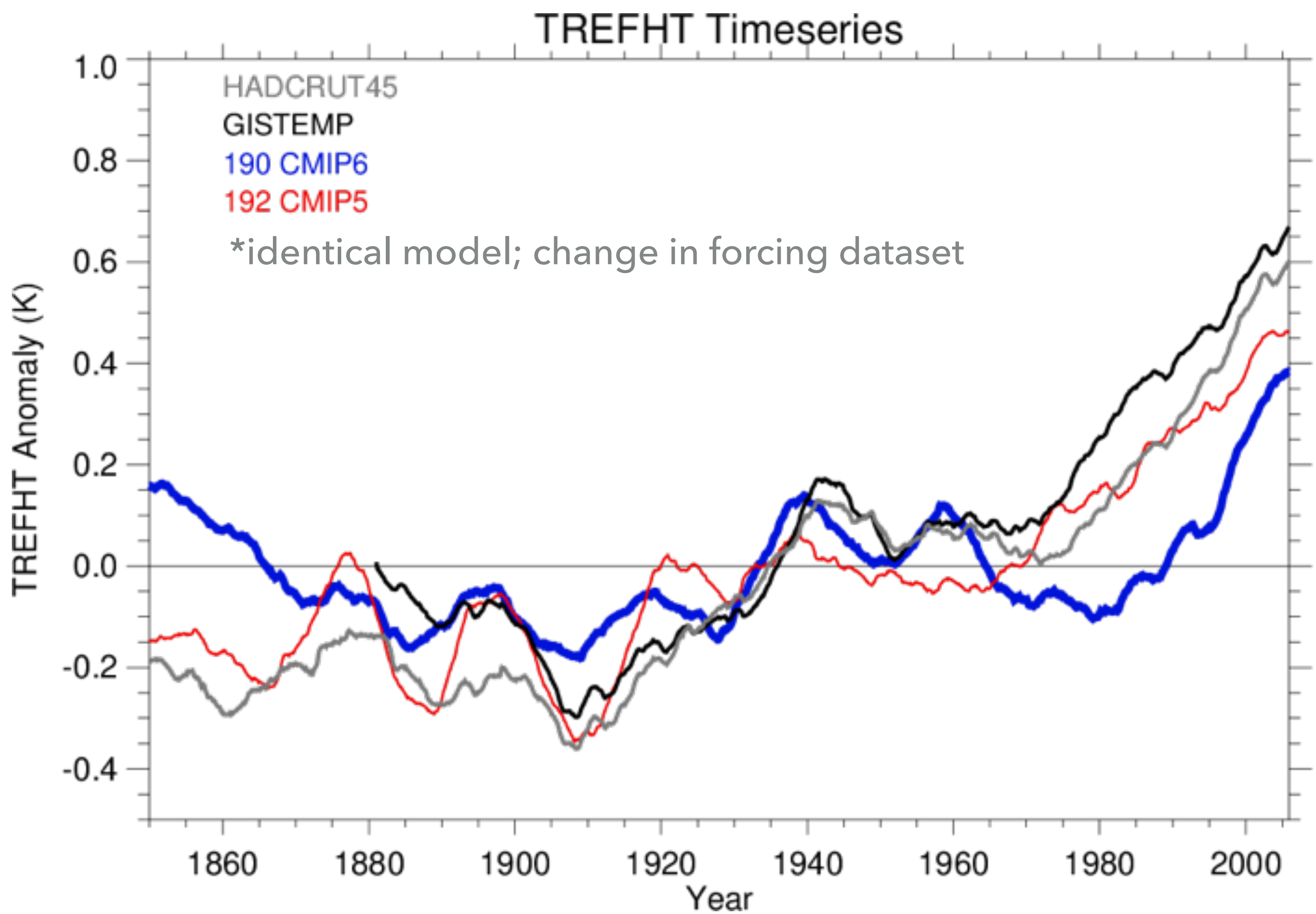
- Challenges involved in climate model evaluation.
- Development of CMAT, an objective and comprehensive model evaluation package.
  - An approach for selecting robust metrics and objectively scoring model development runs (often PI-control).
- Provide context for CMAT scoring of CESM with LENS and CMIP3/5.
- Next Steps for CMAT: Remaining goals and opportunities.
- Revisit CESM2 development versions; discuss CESM2 successes and intriguing behavior.

MOTIVATION: CESM GLOBAL MEAN SURFACE TEMPERATURE



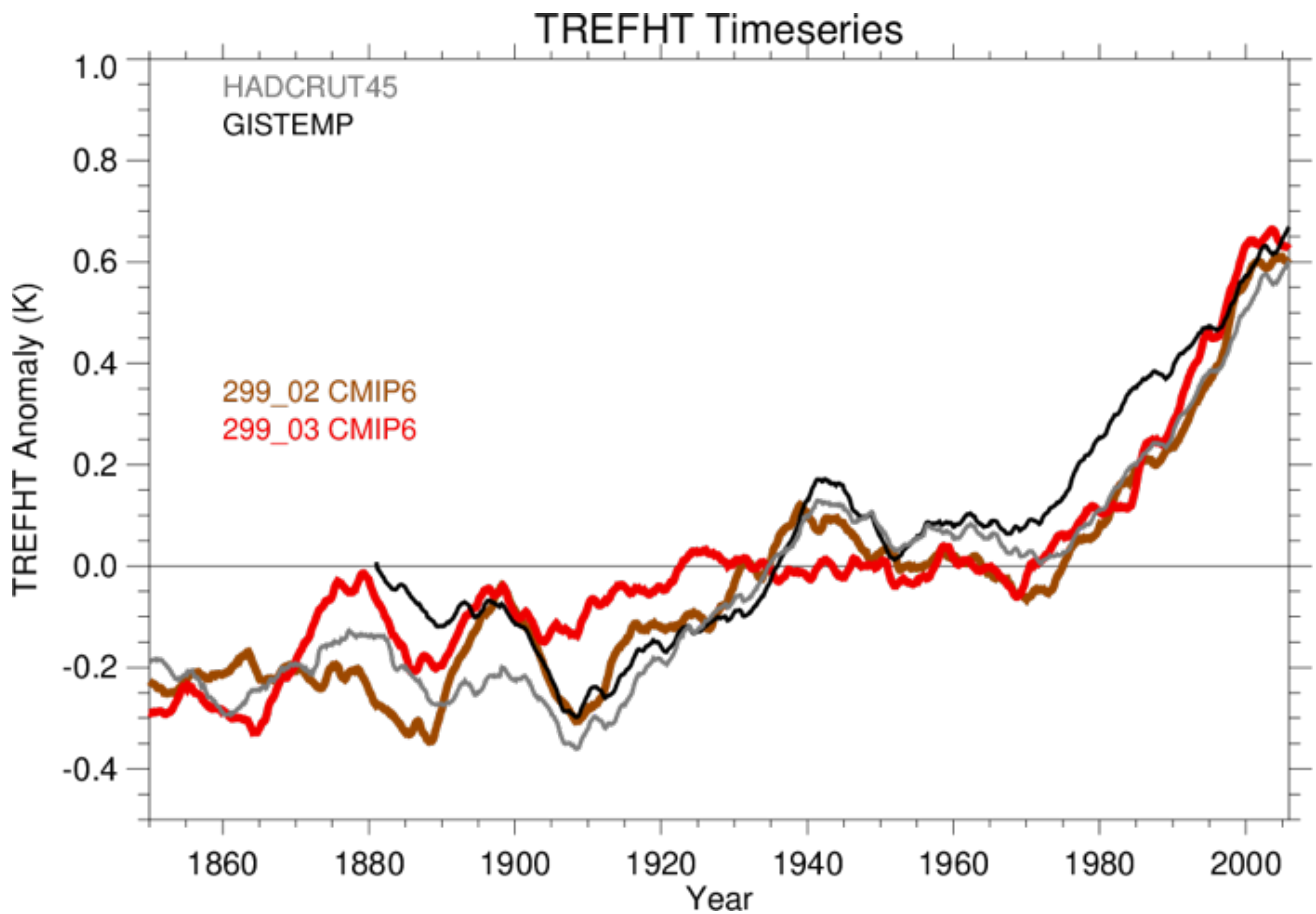
- Early development versions of CESM2 (e.g. 125, circa 2015) warmed as observed…while later versions cooled; both due to changes in the model…

MOTIVATION: CESM GLOBAL MEAN SURFACE TEMPERATURE



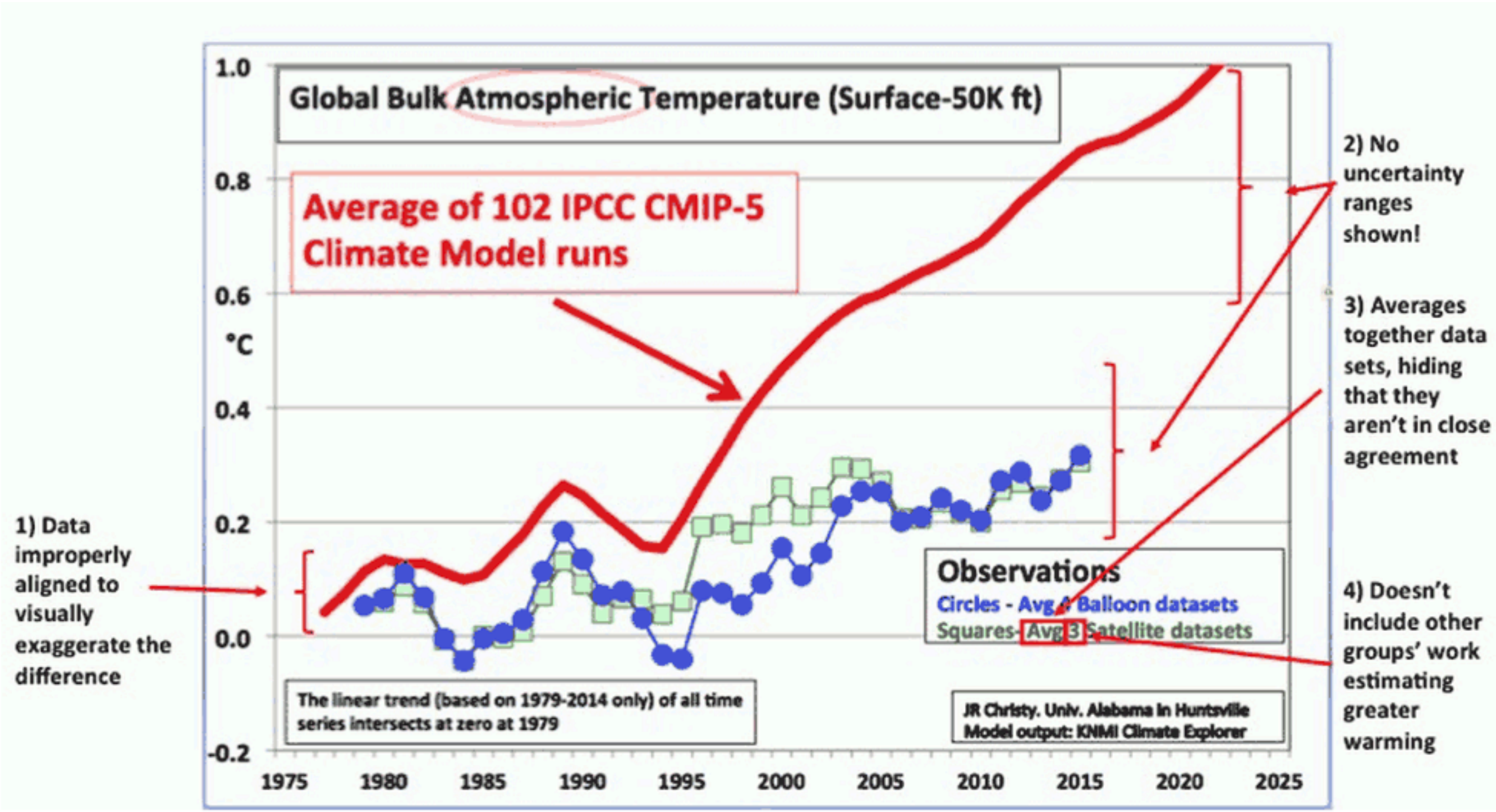
- and due to changes in forcing…190 and 192 are identical model versions forced by CMIP6 and CMIP5 forcing, respectively. Global mean changes in the forcing were largely negligible - patterns were different, particularly for aerosols.

MOTIVATION: CESM GLOBAL MEAN SURFACE TEMPERATURE



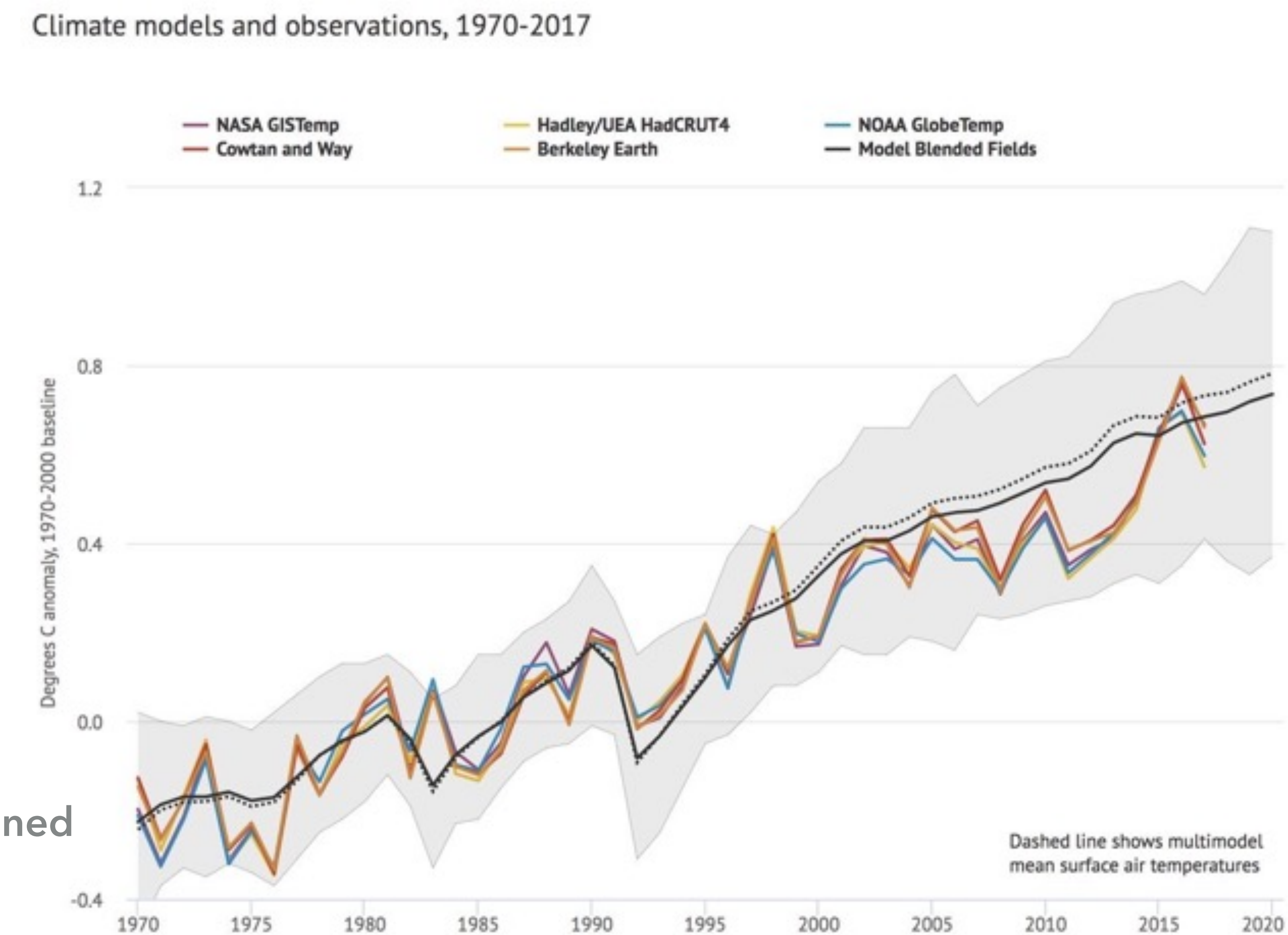
- And in the final CESM2 version (299), we have ensemble members that warm as observed*. *Q: What drove these changes across versions? What changes in the energy budget were involved?*

MODEL EVALUATION: COMMON TECHNICAL PITFALLS



It is also unclear what role uncertainties in forcing play in the differences.
Trends are generally not strong constraints on model fidelity.

MODEL EVALUATION: COMMON TECHNICAL PITFALLS



model spread / internal var. considered

multiple obs best-estimates shown

correctly baselined

But, even when done well, what does it tell you? Due to model feedbacks? Due to errors in forcing? Not a particularly useful diagnostic of model performance - an amalgam of many potential influences. Can't be used to evaluate a PI-control run. Metrics of trends are generally not useful constraints during model development (model drift, internal variability, data uncertainty).

SPECIFIC GOALS OF CMAT

To provide a more informative, comprehensive, and objective evaluation of CESM2 development runs incorporating:

- bounds of internal variability (*when are differences with obs or between model runs meaningful?*)

SPECIFIC GOALS OF CMAT

To provide a more informative, comprehensive, and objective evaluation of CESM2 development runs incorporating:

- bounds of internal variability (*when are differences with obs or between model runs meaningful?*)
- comparison with best available observations *(accounting for uncertainty)* inclusion of physically relevant coupled fields (*e.g. net surface flux, cloud forcing*),

SPECIFIC GOALS OF CMAT

To provide a more informative, comprehensive, and objective evaluation of CESM2 development runs incorporating:

- bounds of internal variability (*when are differences with obs or between model runs meaningful?*)
- comparison with best available observations *(accounting for uncertainty)* inclusion of physically relevant coupled fields (*e.g. net surface flux, cloud forcing*),
- a pre-industrial increment is used when comparing a PI-control run to observations based on CESM1-LENS

SPECIFIC GOALS OF CMAT

To provide a more informative, comprehensive, and objective evaluation of CESM2 development runs incorporating:

- bounds of internal variability (*when are differences with obs or between model runs meaningful?*)
- comparison with best available observations *(accounting for uncertainty)* inclusion of physically relevant coupled fields (*e.g. net surface flux, cloud forcing*),
- a pre-industrial increment is used when comparing a PI-control run to observations based on CESM1-LENS
- quantification of model drift based on ocean heat content at many levels

SPECIFIC GOALS OF CMAT

To provide a more informative, comprehensive, and objective evaluation of CESM2 development runs incorporating:

- bounds of internal variability (*when are differences with obs or between model runs meaningful?*)
- comparison with best available observations *(accounting for uncertainty)* inclusion of physically relevant coupled fields (*e.g. net surface flux, cloud forcing*),
- a pre-industrial increment is used when comparing a PI-control run to observations based on CESM1-LENS
- quantification of model drift based on ocean heat content at many levels
- context vs CMIP archives (*is the model in question a 'good' model?*)

SPECIFIC GOALS OF CMAT

To provide a more informative, comprehensive, and objective evaluation of CESM2 development runs incorporating:

- bounds of internal variability (*when are differences with obs or between model runs meaningful?*)
- comparison with best available observations *(accounting for uncertainty)* inclusion of physically relevant coupled fields (*e.g. net surface flux, cloud forcing*),
- a pre-industrial increment is used when comparing a PI-control run to observations based on CESM1-LENS
- quantification of model drift based on ocean heat content at many levels
- context vs CMIP archives (*is the model in question a 'good' model?*)
- minimal redundancy with existing diagnostics packages (*e.g. CVDP, AMWG*).

SPECIFIC GOALS OF CMAT

To provide a more informative, comprehensive, and objective evaluation of CESM2 development runs incorporating:

- bounds of internal variability (*when are differences with obs or between model runs meaningful?*)
- comparison with best available observations *(accounting for uncertainty)* inclusion of physically relevant coupled fields (*e.g. net surface flux, cloud forcing*),
- a pre-industrial increment is used when comparing a PI-control run to observations based on CESM1-LENS
- quantification of model drift based on ocean heat content at many levels
- context vs CMIP archives (*is the model in question a 'good' model?*)
- minimal redundancy with existing diagnostics packages (*e.g. CVDP, AMWG*).
- a summary resource for understanding the relative performance across many runs

MODEL EVALUATION: METRICS

EOS Earth & Space Science News

ATMOSPHERIC SCIENCES Opinion

Human-Caused Warming or Just Natural Variations?

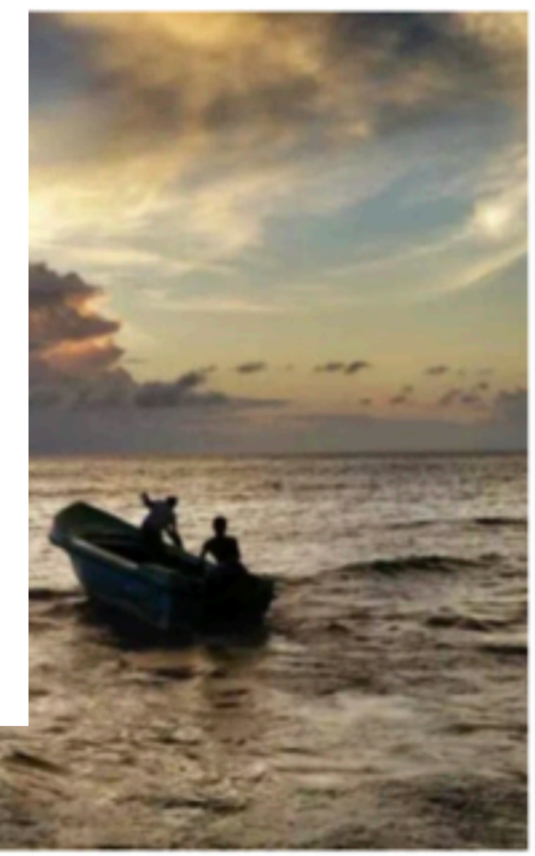
Planet

content and sea level rise
ible answer than atmospheric

The amplitude of the global warming signature (signal) compared with natural variability (noise) defines how well a metric tracks global warming. The "noise level," that is, the amplitude of internal variability, approximated here by the standard deviation (σ) of the OHC time series after the linear trend is removed, amounts to 0.77×10^{22} J from 2004 to 2015 (Table 1). The linear trend of OHC is $0.79 \pm 0.03 \times 10^{22}$ J year^{-1} within the same period (Figure 2).

Table 1. The Linear Trend (with 95% Confidence Level) for the Three Key Climate Indicators: Global Mean Surface Temperature (GMST), Ocean Heat Content (OHC), and Sea Level Rise (SLR)[a]

	Linear Trend	σ	S/N (1/years)	Time (years)
GMST	0.016°C ± 0.005°C/yr	0.110°C/yr	0.14	27
OHC	$0.79 \pm 0.03 \times 10^{22}$ J/yr	0.77×10^{22} J/yr	1.03	3.9
SLR	3.38 ± 0.10 mm/yr	3.90 mm/yr	0.87	4.6



Fishermen ply their trade in the Gulf of Mannar in the Indian Ocean near Sri Lanka, one example of the vital importance of oceans to planet Earth and humankind. Measurements of ocean heating and sea level rise could prove more reliable than atmospheric measurements for tracking vital signs for the health of the planet. Credit: Jiang Zhu

By Lijing Cheng, Kevin E. Trenberth, John Fasullo, John Abraham, Tim P. Boyer, Karina von Schuckmann, and Jiang Zhu 13 September 2017

MODEL EVALUATION: METRICS

Q: Why focus on the energy budget?

Burrows et al., 2018: AAS

Survey of 96 (62>10yrs, 31>20yrs experience) climate scientists.

- results from a large community survey on the relative importance of different variables in evaluating a climate model's fidelity
- no statistically significant differences between rankings provided by model developers and model users
- "high" and "low" experience groups were in agreement about the importance of most variables
- limited the scope of the study to evaluation of global mean climate but recognized the need to extend this (diurnal/AC/ENSO…)

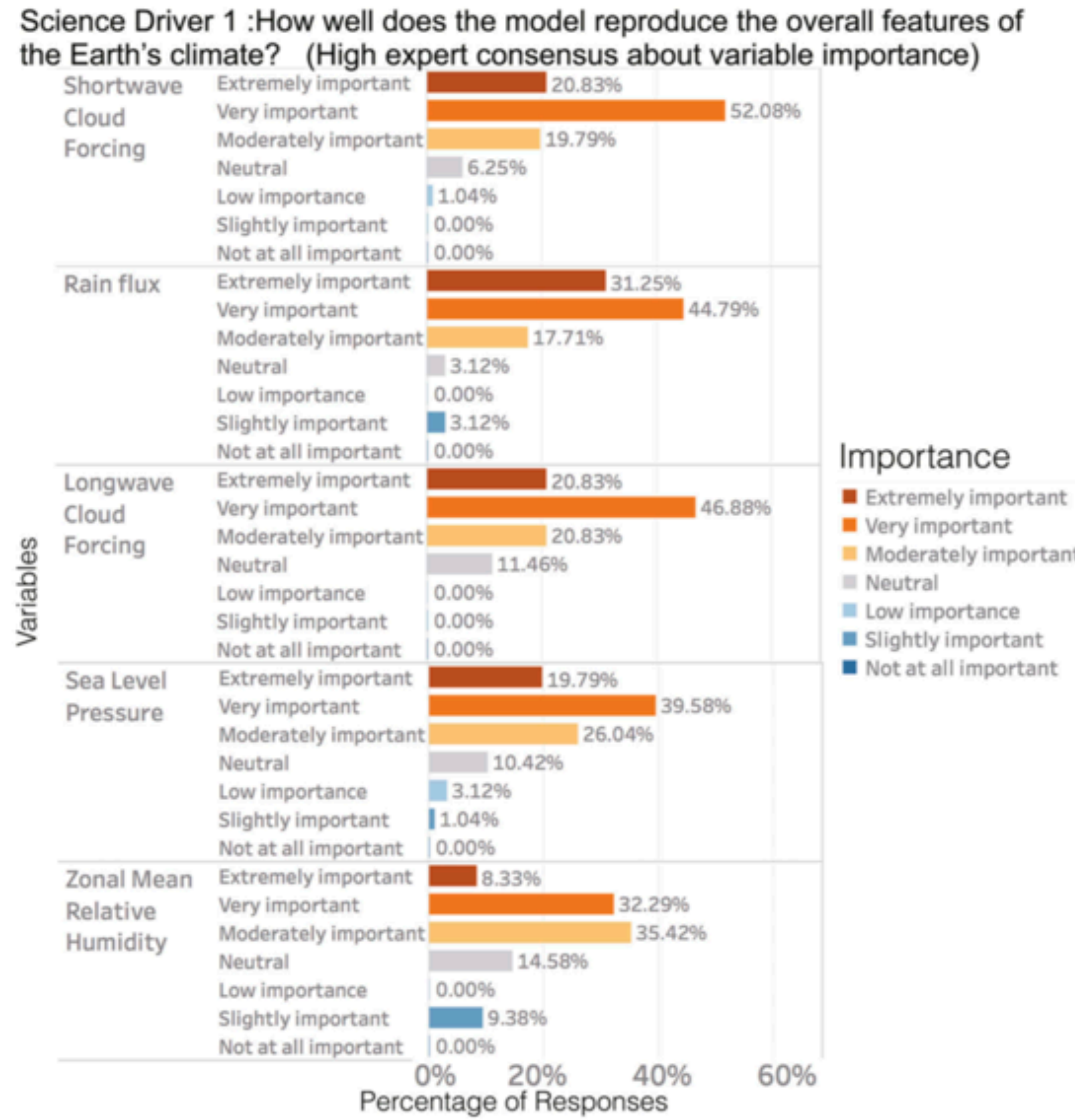
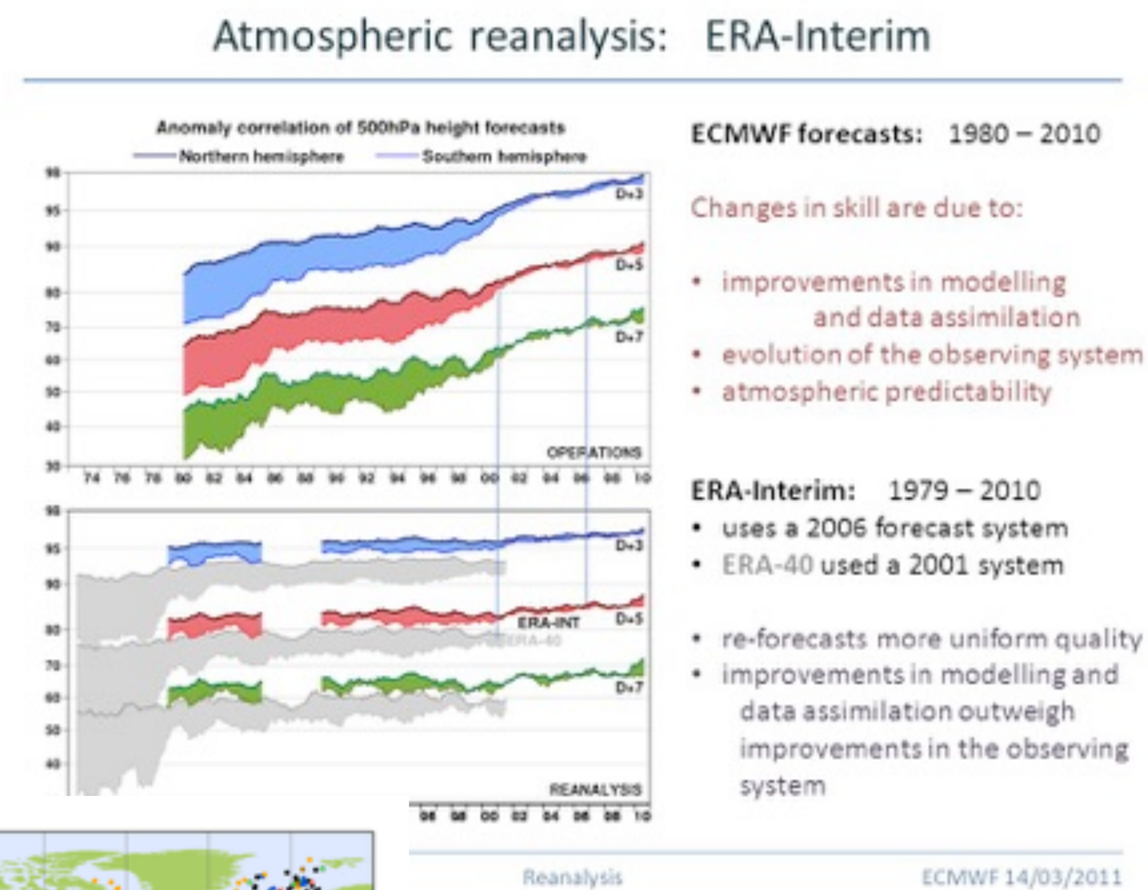


Fig. 2. Science Driver 1: distributions of importance ratings, ranked by consensus, as quantified by the coefficient of agreement *A*, for variables with high expert consensus about their importance.

MODEL EVALUATION: OBSERVATIONAL ADVANCES

JPSS



ARGO

MODEL EVALUATION: MODELING ADVANCES: LENS

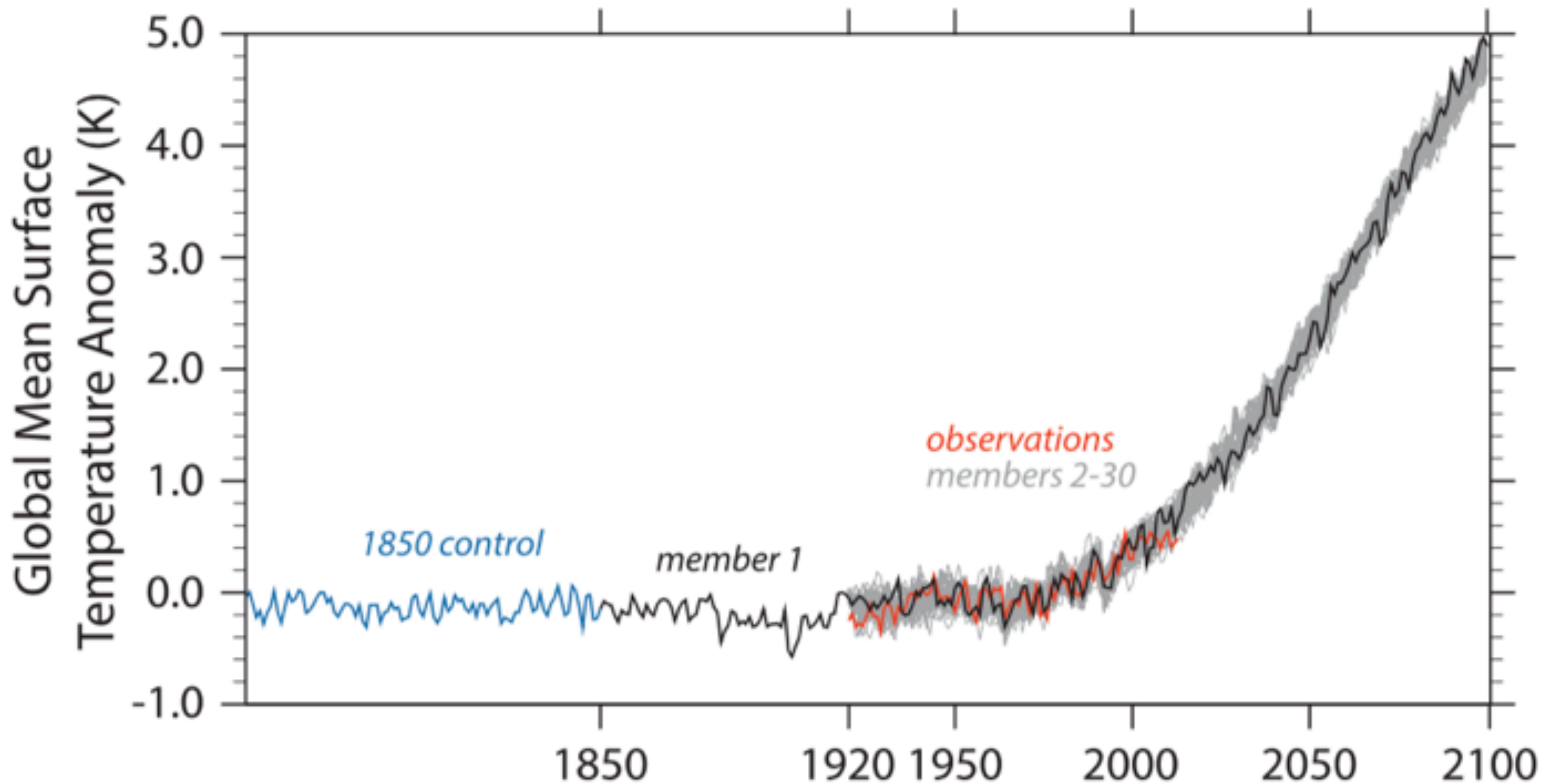


Fig. 2. Global surface temperature anomaly (1961–90 base period) for the 1850 control, individual ensemble members, and observations (HadCRUT4; Morice et al. 2012).

The CESM1 Large Ensemble provides estimates of:

1) the range of variability that is internal to the climate system (significance) and
2) the magnitude and patterns of the forced climate response (pi-increment).

(caveat - both can be model dependent)

CESM2 AND CMAT

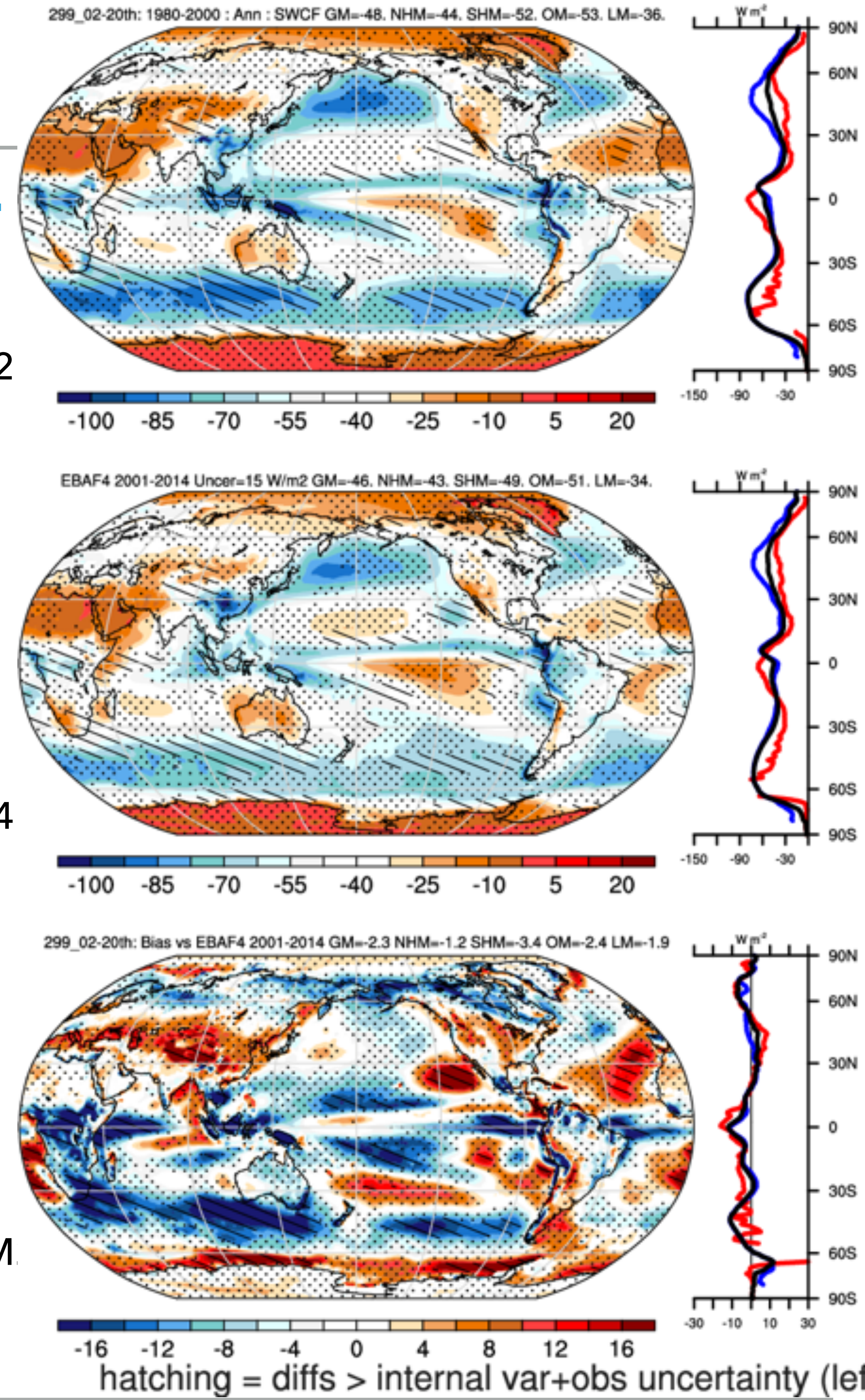
METRIC #1: CLIMATOLOGICAL MEAN: SWCF

- Annual mean for 20 yrs (1980-2000) climatologies are compiled. Zonal means on right for land, ocean and **both**.
- If the model runs is PI, the LENS increment is used to account for the difference. (this example uses an historical run).
- **Stippling** (same in all panels) is applied where differences are larger than can be explained by internal variability.
- **Hatching** (also same) is applied where differences are larger than can be explained by internal variability plus observational uncertainty.
- CESM2 SWCF biases are ~10 Wm^{-2}; less on a zonal mean basis. This is a major advance provided in part by CLUBB.

CESM2

EBAF4

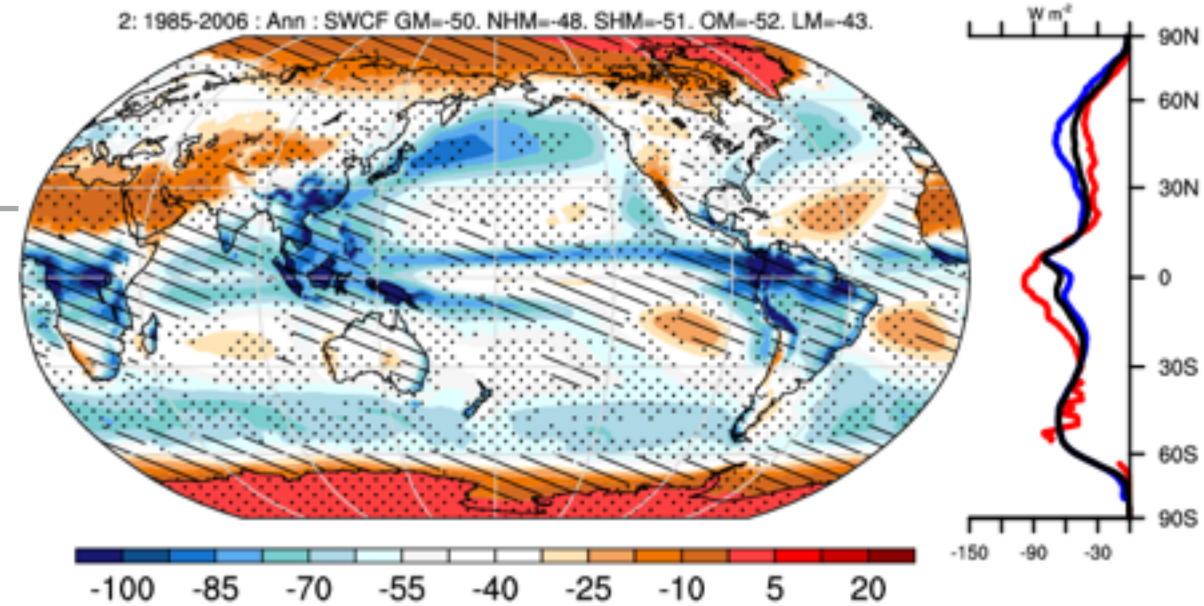
CESM BIAS



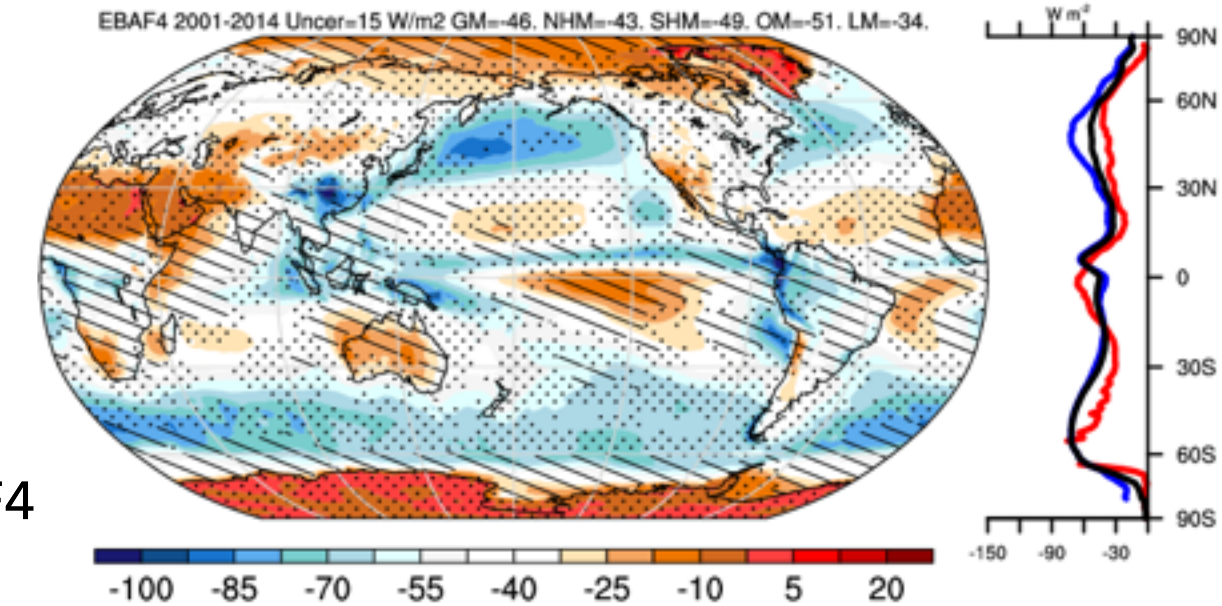
METRIC #1: CLIMATOLOGICAL MEAN: SWCF

- CESM2 is a major advance over earlier models.
- SWCF biases were substantially larger in earlier climate models. CESM1 (LENS) biases are substantially larger both regionally and on a zonal mean basis (~20 Wm^{-2})
- CESM1 tended to overcool the tropics, undercool high latitudes. CESM2 moves significant additional SWCF to high latitudes and therefore increases the associated transfers of energy. (Less SW into tropical regions; more SW into polar regions.)

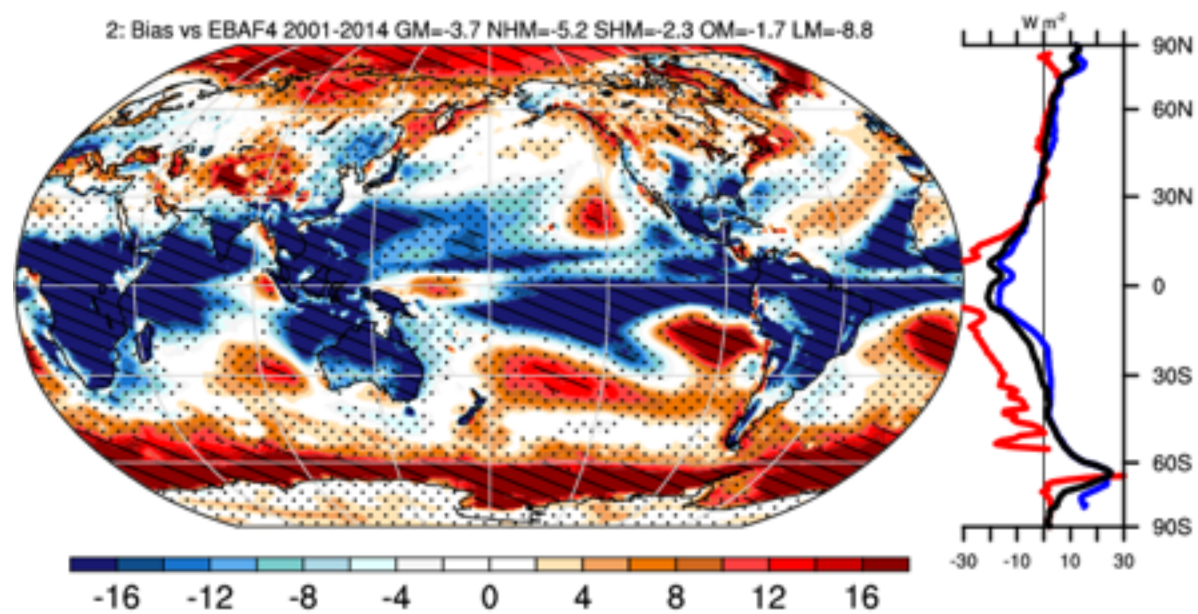
LENS



EBAF4



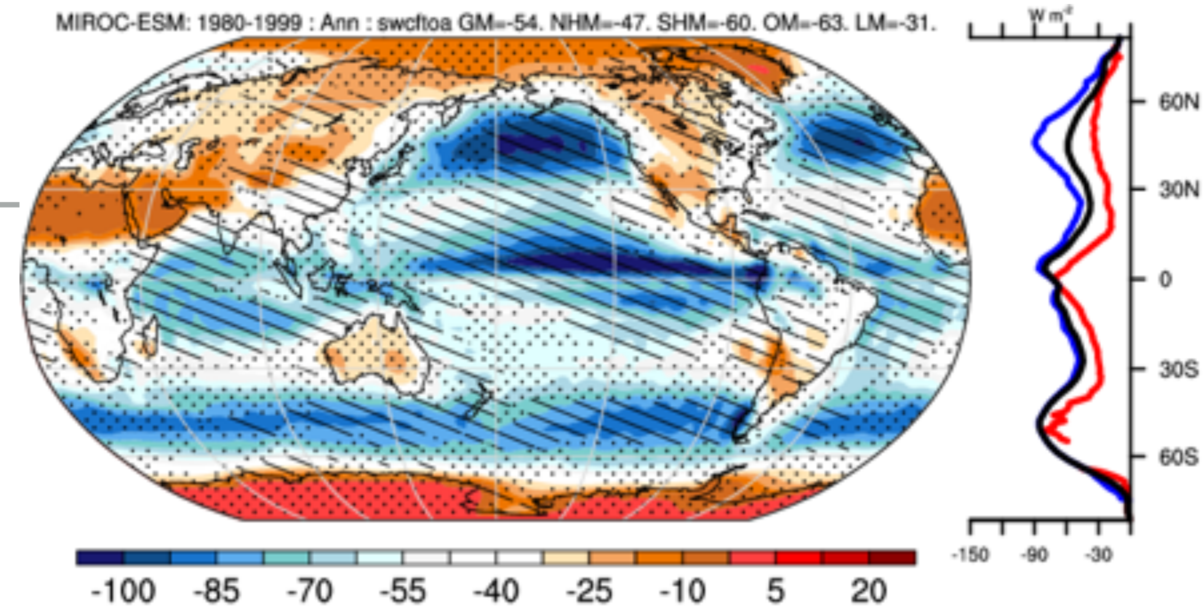
LENS BIAS



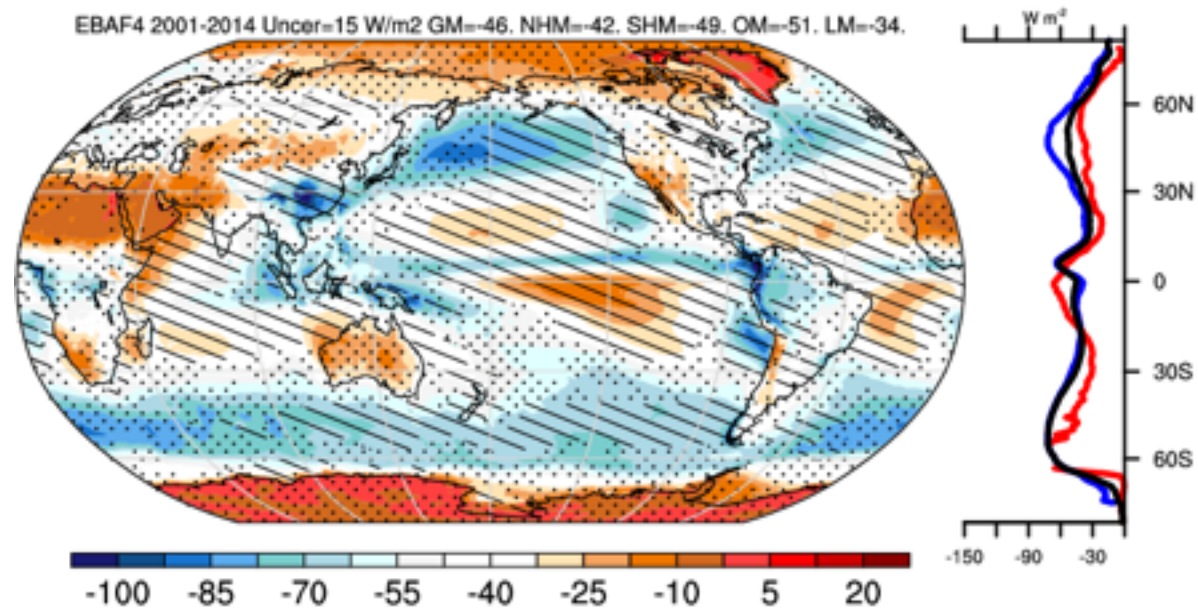
hatching = diffs > internal var+obs uncertainty (lef

METRIC #1: CLIMATOLOGICAL MEAN: SWCF

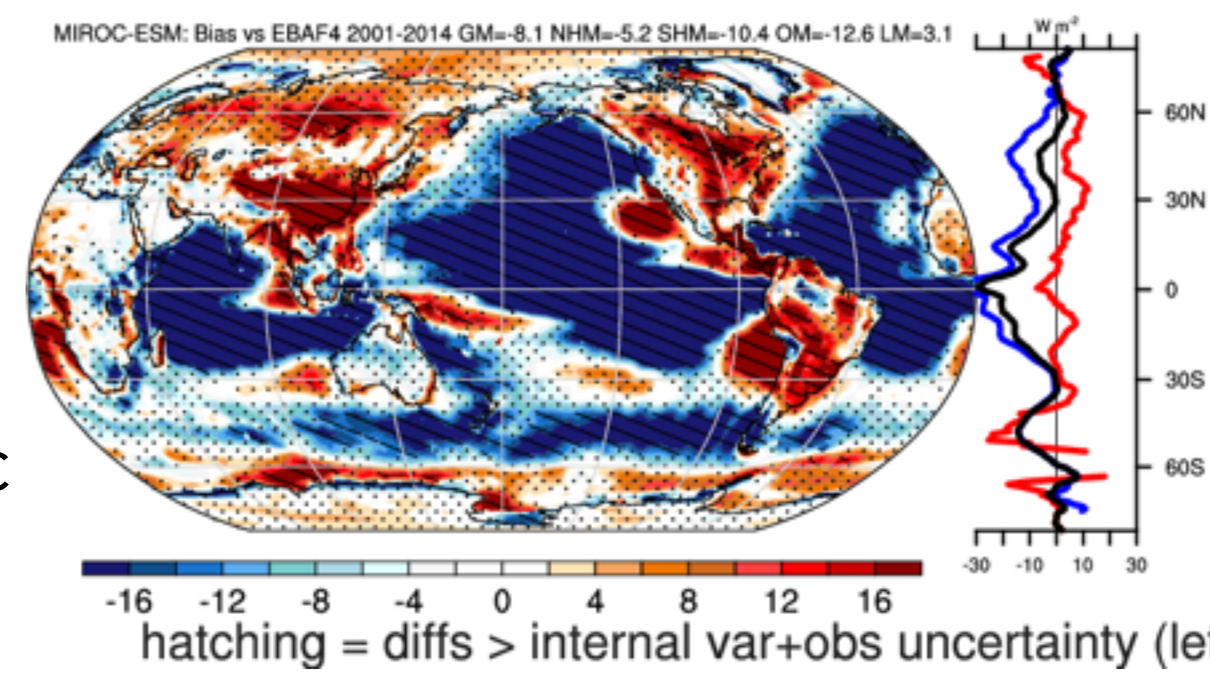
MIROC



EBAF4



MIROC BIAS

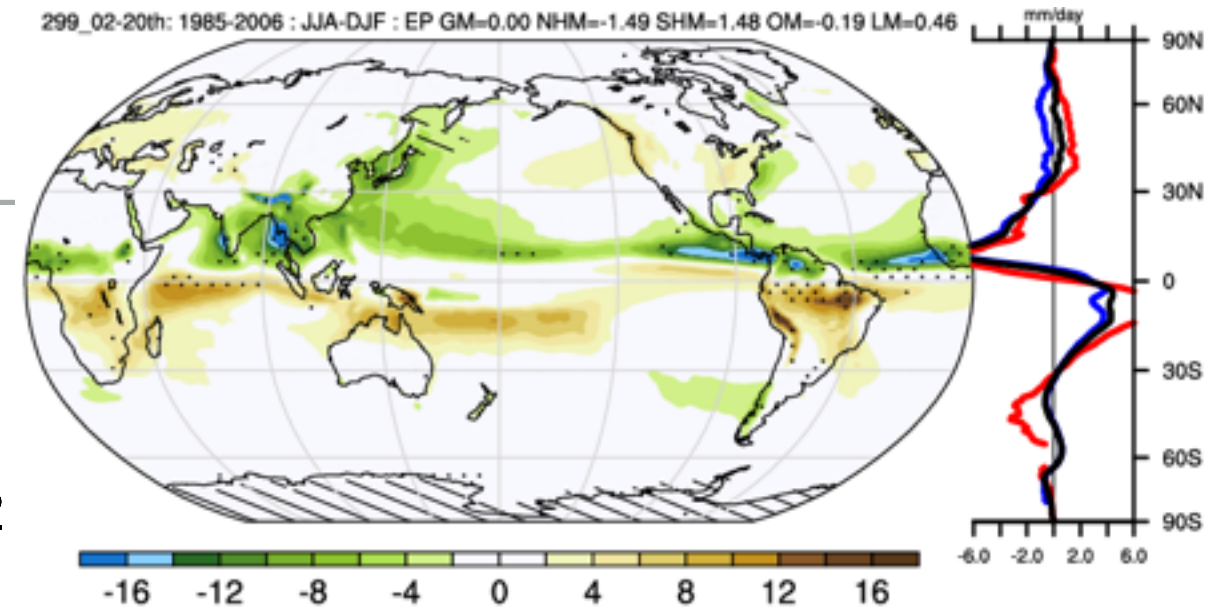


- In other CMIP5 models SWCF biases were even larger than CESM1 and despite not being too bad in the global mean (biases ~10 Wm^{-2}) exhibited substantial regional (>30 Wm^{-2}) and zonal mean (~20-30 Wm^{-2}) biases.
- Biases are again systematic with latitude with too little SW flux into tropical regions and too much into polar regions ➔ fundamentally different flows of energy in the climate system.

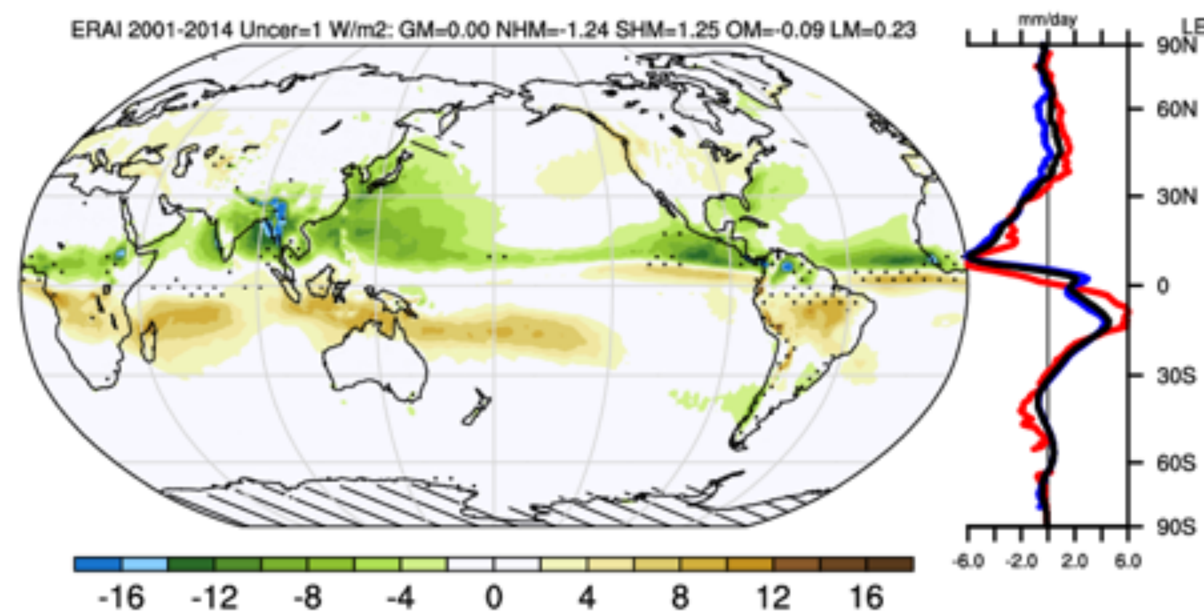
METRIC #2: SEASONAL CONTRASTS: JJA-DJF : E-P

- One approach for assessing seasonality is with seasonal or monthly means - but this involves significant redundancy with climatological mean assessment (metric #1).
- Instead seasonal differences (JJA-DJF) can be used to better isolate seasonal variability and be scored through pattern correlations.
- E-P - in CMAT estimated from the atmospheric moisture budget of ERA Interim from the analysis not forecast fields - is an example of a field with significant seasonality that remains a challenge for models to capture. The pattern correlation for CESM2 is only marginally better than CESM1.

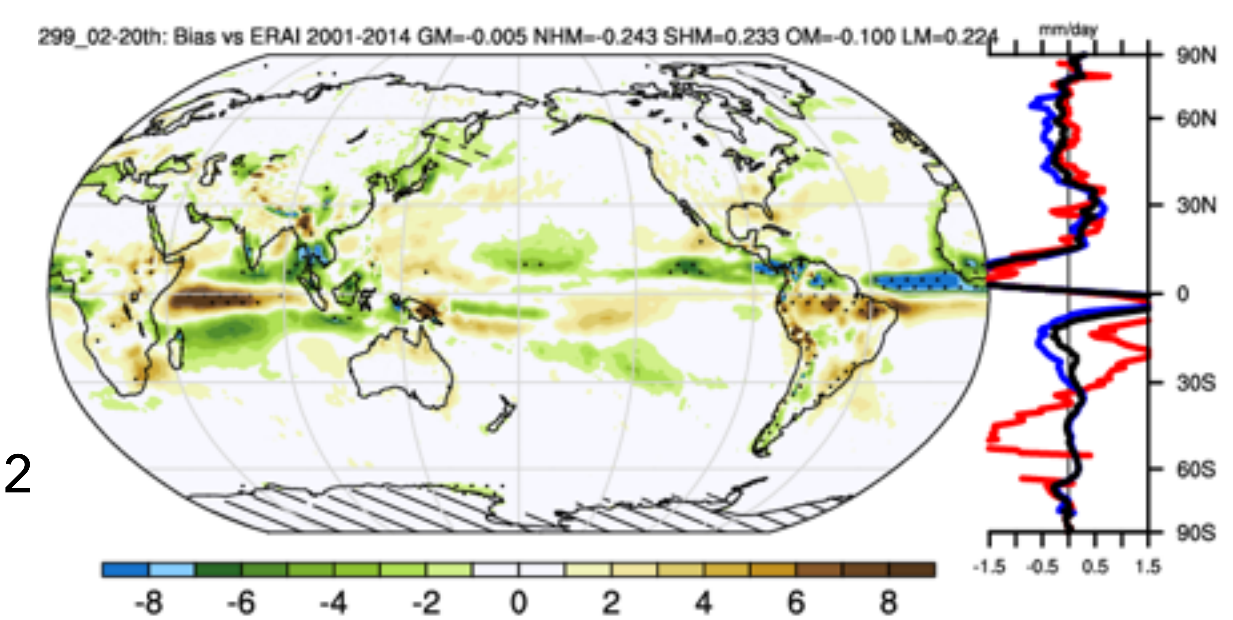
CESM2



ERA-I



CESM2 BIAS

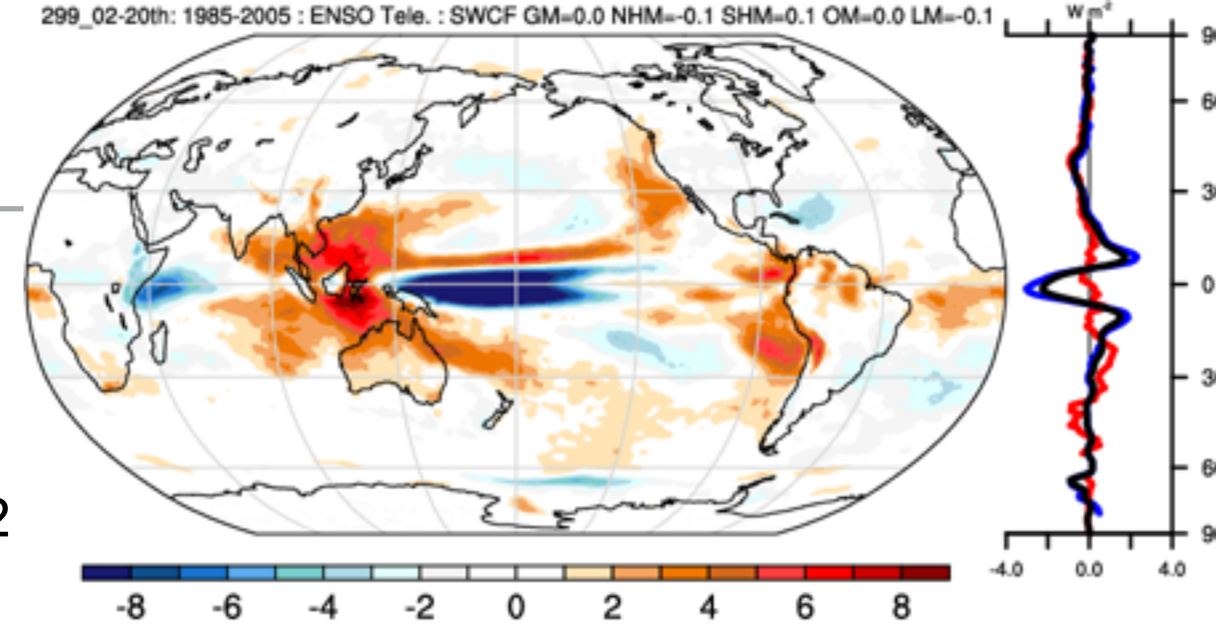


hatching = diffs > internal var+obs uncertainty (le

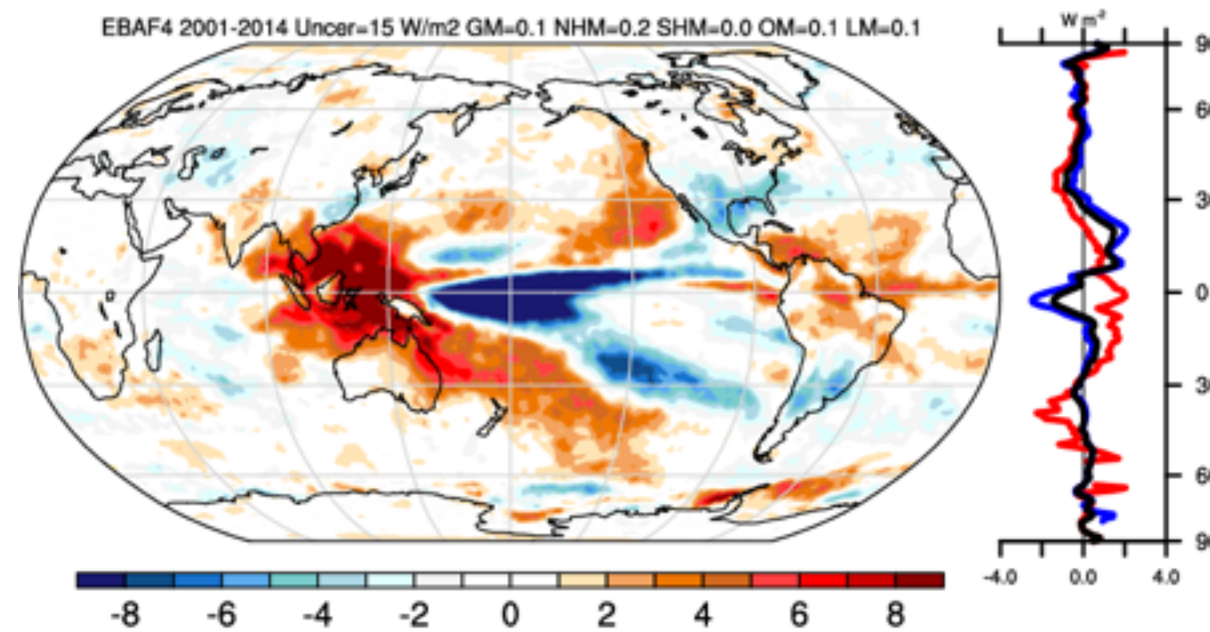
METRIC #3: ENSO PATTERNS : SWCF

- ENSO is often assessed with seasonal means (e.g. DJF).
- A challenge exists however in having a record long enough to provide a meaningful measure of ENSO's spatial patterns.
- As a more robust measure, the July through June regression of Niño3.4 SST anomalies with remote anomalies (in phase with ENSO events) is used - provides a more robust measure than DJF alone.

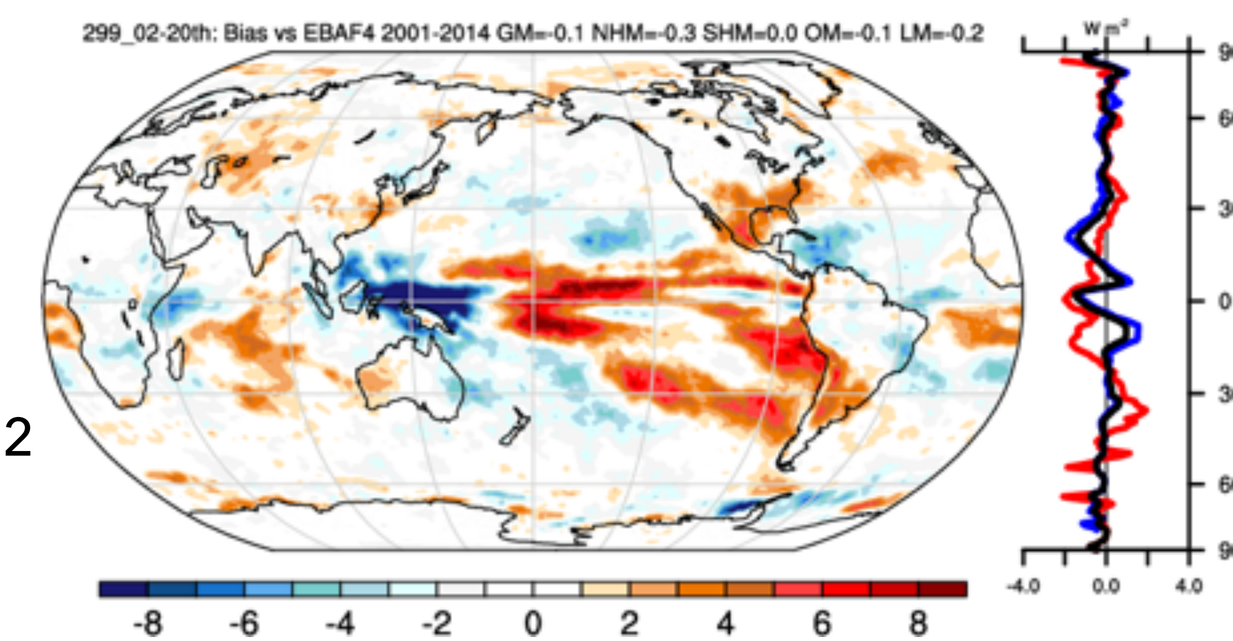
CESM2



CERES



CESM2
BIAS



MODEL SCORING:

- **To simplify comparison across model versions, Variable and Overall Scores are computed.**

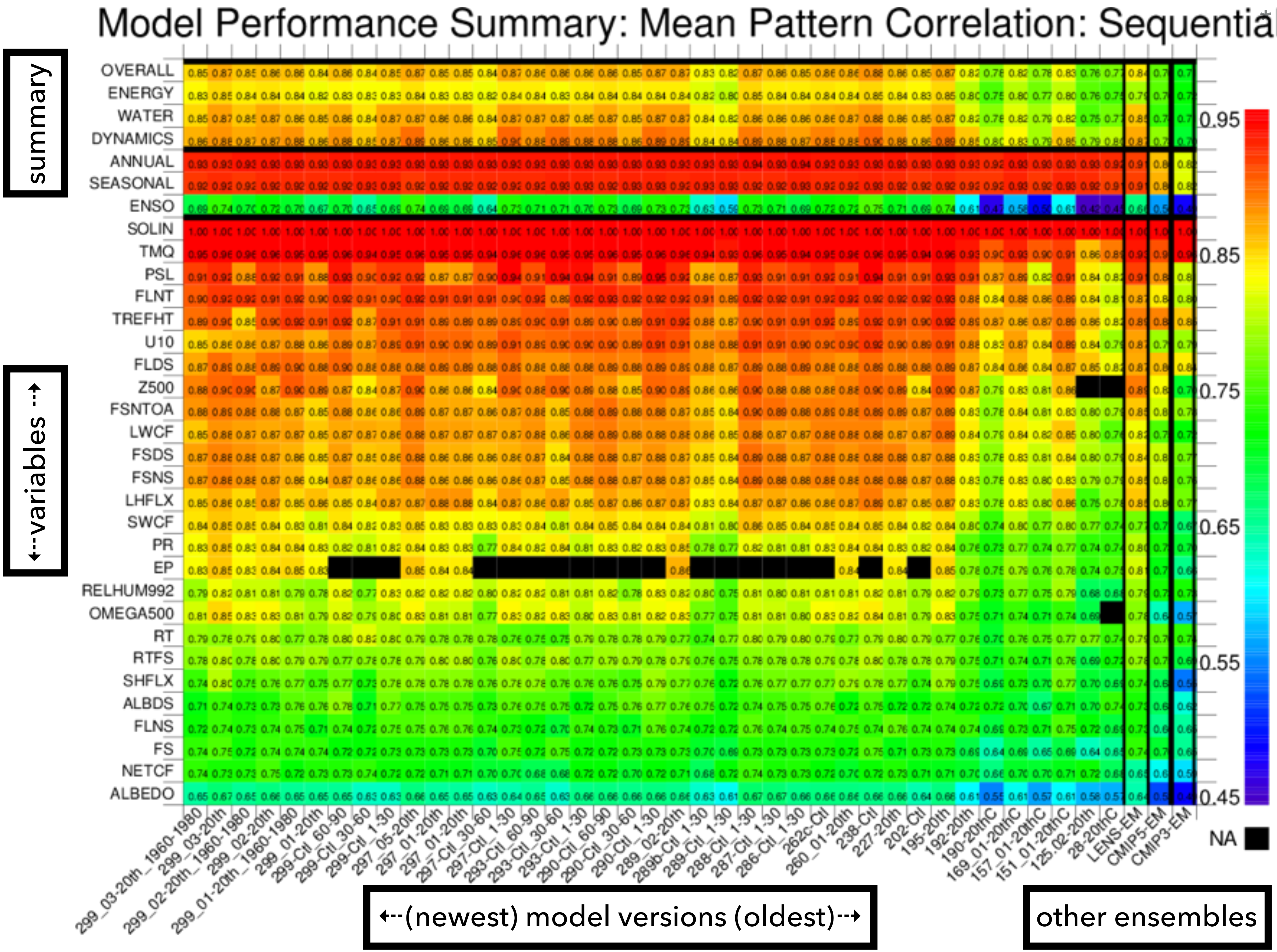
- **Variable scores:** computed from the weighted average of 3 pattern correlations (mean, seasonal, and ENSO metrics) using:

  I. **Energy Budget** (FSNT, FLNT, SWCF, LWCF, F_S, and R_T-F_S)

  II. **Water Cycle** (PRECTOT, TMQ, RELHUM995, LH, EP)

  III. **Dynamics** (PSL, U10, Z500, OMEGA500)

- **Overall Score**: the average of the energy budget, water cycle, and dynamics scores. Metric weights for variable scores are set such that the 1-σ range due to internal variability in the overall score is 0.01 (based on CESM1-LENS).


Model Performance Summary: Mean Pattern Correlation: Sequential
summary
←--variables --→
OVERALL
ENERGY
WATER
DYNAMICS
ANNUAL
SEASONAL
ENSO
SOLIN
TMQ
PSL
FLNT
TREFHT
U10
FLDS
Z500
FSNTOA
LWCF
FSDS
FSNS
LHFLX
SWCF
PR
EP
RELHUM992
OMEGA500
RT
RTFS
SHFLX
ALBDS
FLNS
FS
NETCF
ALBEDO
299_03-20th_1960-1980
299_03-20th
299_02-20th_1960-1980
299_02-20th
299_01-20th_1960-1980
299_01-20th
299-Ctl_60-90
299-Ctl_30-60
299-Ctl_1-30
297_05-20th
297_01-20th
297_01-20th
297-Ctl_30-60
297-Ctl_1-30
293-Ctl_60-90
293-Ctl_30-60
293-Ctl_1-30
290-Ctl_60-90
290-Ctl_30-60
290-Ctl_1-30
289_02-20th
289b-Ctl_1-30
289-Ctl_1-30
288-Ctl_1-30
287-Ctl_1-30
286-Ctl_1-30
262c-Ctl
260_01-20th
238-Ctl
227-20th
202-Ctl
195-20th
192-20th
190-20thC
169_01-20thC
157_01-20thC
151_01-20thC
125.02-20th
28-20thC
LENS-EM
CMIP5-EM
CMIP3-EM
0.95
0.85
0.75
0.65
0.55
0.45
BETTER
NA
←--(newest) model versions (oldest)--→
other ensembles

CMAT SINGLE RUN PAGES

- Run pages also provide numerous additional fields for diagnosing model behavior including:
  - global and regional time series
  - Hovmöller diagrams,
  - summary statements on improvements, degradations, and flags identifying other biases.

NCAR UCAR | CGD's Climate Analysis Section's
Climate Model Assessment Tool

Methodology
CMAT Repository
Created: Mon Oct 2 15:25:13 2017
CMAT Version 0.1

CMAT 0.1 CESM 192-20th: vs 190-20th:
(Overall: 0.83-A; Energy: 0.80-A; Water: 0.82-A; Dynamics: 0.89-A)

Summary Comments

Notes	Flags vs Obs	Improvements (vs 190-20th)	Degradations (vs 190-20th)
same as 190+ cmip5 emissions + start from 190 at year 64 instead of LENS at yr 402	Global Mean FSNTOA too small, see also SWCF, ALBDS Global Mean FLNT too small, see also LWCF, TREFHT, TMQ OM-LM LHFLX too small.	FSNTOA PCor-ENSO improves, see also SWCF, ALBDS FLNT PCor-ENSO improves, see also LWCF, TREFHT, TMQ SWCF PCor-ENSO improves, see also ALBDS LWCF PCor-ENSO improves, FS PCor-ENSO improves, LHFLX PCor-ENSO improves, RTFS PCor-ENSO improves, Z500 PCor-ENSO improves, RELHUM992 PCor-ENSO improves, U10 PCor-ENSO improves, PSL PCor-ENSO improves,	

Variable Options

Regional Options

Select Timeseries

Global: TREFHT, RT, FSNTOA, FLNT, SWCF, LWCF, OHC 0-300m, OHC 0-700m, OHC 0-2000, P, CLDTOT, CLDLOW, CLDHGH, TMQ

NH: TREFHT, RT, FSNTOA, FLNT, SWCF, LWCF, OHC 0-300m, OHC 0-700m, OHC 0-2000, P, CLDTOT, CLDLOW, CLDHGH, TMQ

SH: TREFHT, RT, FSNTOA, FLNT, SWCF, LWCF, OHC 0-300m, OHC 0-700m, OHC 0-2000, P, CLDTOT, CLDLOW, CLDHGH, TMQ

Ocean: TREFHT, RT, FSNTOA, FLNT, SWCF, LWCF, P, CLDTOT, CLDLOW, CLDHGH, TMQ

Land: TREFHT, RT, FSNTOA, FLNT, SWCF, LWCF, P, CLDTOT, CLDLOW, CLDHGH, TMQ

Tropics(30/30): TREFHT, RT, FSNTOA, FLNT, SWCF, LWCF, OHC 0-300m, OHC 0-700m, OHC 0-2000, P, CLDTOT, CLDLOW, CLDHGH, TMQ

Ocean(30/30): TREFHT, RT, FSNTOA, FLNT, SWCF, LWCF, P, CLDTOT, CLDLOW, CLDHGH, TMQ

Land(30/30): TREFHT, RT, FSNTOA, FLNT, SWCF, LWCF, P, CLDTOT, CLDLOW, CLDHGH, TMQ

Global (by LME member): TREFHT, FSNTOA, FLNT, SWCF, LWCF, OHC 0-300m, OHC 0-700m, OHC 0-2000, P, CLDTOT, CLDLOW, CLDHGH, TMQ

Ocean Drift: Summary **Nino3.4:** TS **Forcing Est.** FSNTC 30/30 Ocean

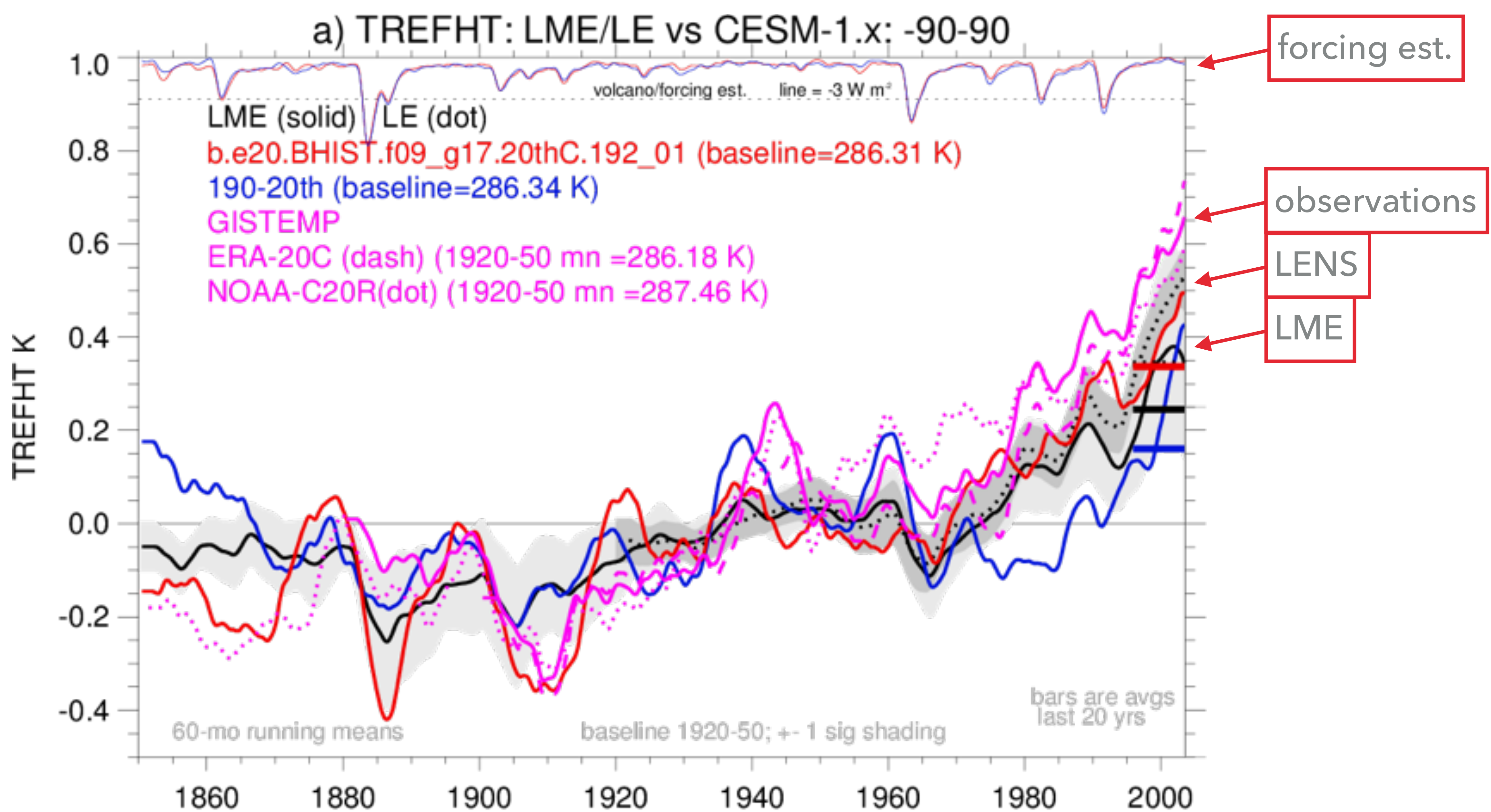
Lat/Lon Maps

HOVs

Lat/Lon Maps

Key Energy Variables	*Scored/Graded*			*Unscored*			
FSNTOA 0.83	Annual 0.99	JJA-DJF 0.99	ENSO Tele 0.51	DJF 0.99	MAM 0.99	JJA 0.99	SON 0.99
FLNT 0.88	Annual 0.98	JJA-DJF 0.95	ENSO Tele 0.71	DJF 0.97	MAM 0.96	JJA 0.98	SON 0.97
SWCF 0.80	Annual 0.91	JJA-DJF 0.94	ENSO Tele 0.53	DJF 0.94	MAM 0.87	JJA 0.92	SON 0.90
LWCF 0.83	Annual 0.92	JJA-DJF 0.90	ENSO Tele 0.67	DJF 0.91	MAM 0.86	JJA 0.93	SON 0.90
FS 0.68	Annual 0.79	JJA-DJF 0.97	ENSO Tele 0.29	DJF 0.95	MAM 0.86	JJA 0.93	SON 0.87
RTFS 0.75	Annual 0.92	JJA-DJF 0.91	ENSO Tele 0.42	DJF 0.93	MAM 0.89	JJA 0.89	SON 0.92
Key Water Variables	*Scored/Graded*			*Unscored*			
PR 0.75	Annual 0.87	JJA-DJF 0.84	ENSO Tele 0.55	DJF 0.86	MAM 0.84	JJA 0.84	SON 0.84
TMQ 0.93	Annual 0.99	JJA-DJF 0.97	ENSO Tele 0.83	DJF 0.99	MAM 0.99	JJA 0.99	SON 0.99
LHFLX 0.83	Annual 0.97	JJA-DJF 0.91	ENSO Tele 0.61	DJF 0.96	MAM 0.96	JJA 0.95	SON 0.95
EP 0.78	Annual 0.89	JJA-DJF 0.85	ENSO Tele 0.59	DJF 0.86	MAM 0.85	JJA 0.87	SON 0.86
RELHUM992 0.79	Annual 0.93	JJA-DJF 0.88	ENSO Tele 0.56	DJF 0.90	MAM 0.92	JJA 0.92	SON 0.92
Key Dynamics Variables	*Scored/Graded*			*Unscored*			
PSL 0.91	Annual 0.99	JJA-DJF 0.94	ENSO Tele 0.80	DJF 0.99	MAM 0.98	JJA 0.98	SON 0.99
U10 0.88	Annual 0.98	JJA-DJF 0.94	ENSO Tele 0.72	DJF 0.97	MAM 0.97	JJA 0.97	SON 0.97
Z500 0.87	Annual 0.93	JJA-DJF 0.91	ENSO Tele 0.76	DJF 0.96	MAM 0.87	JJA 0.86	SON 0.91

SUPPORTING METRICS: TIME SERIES



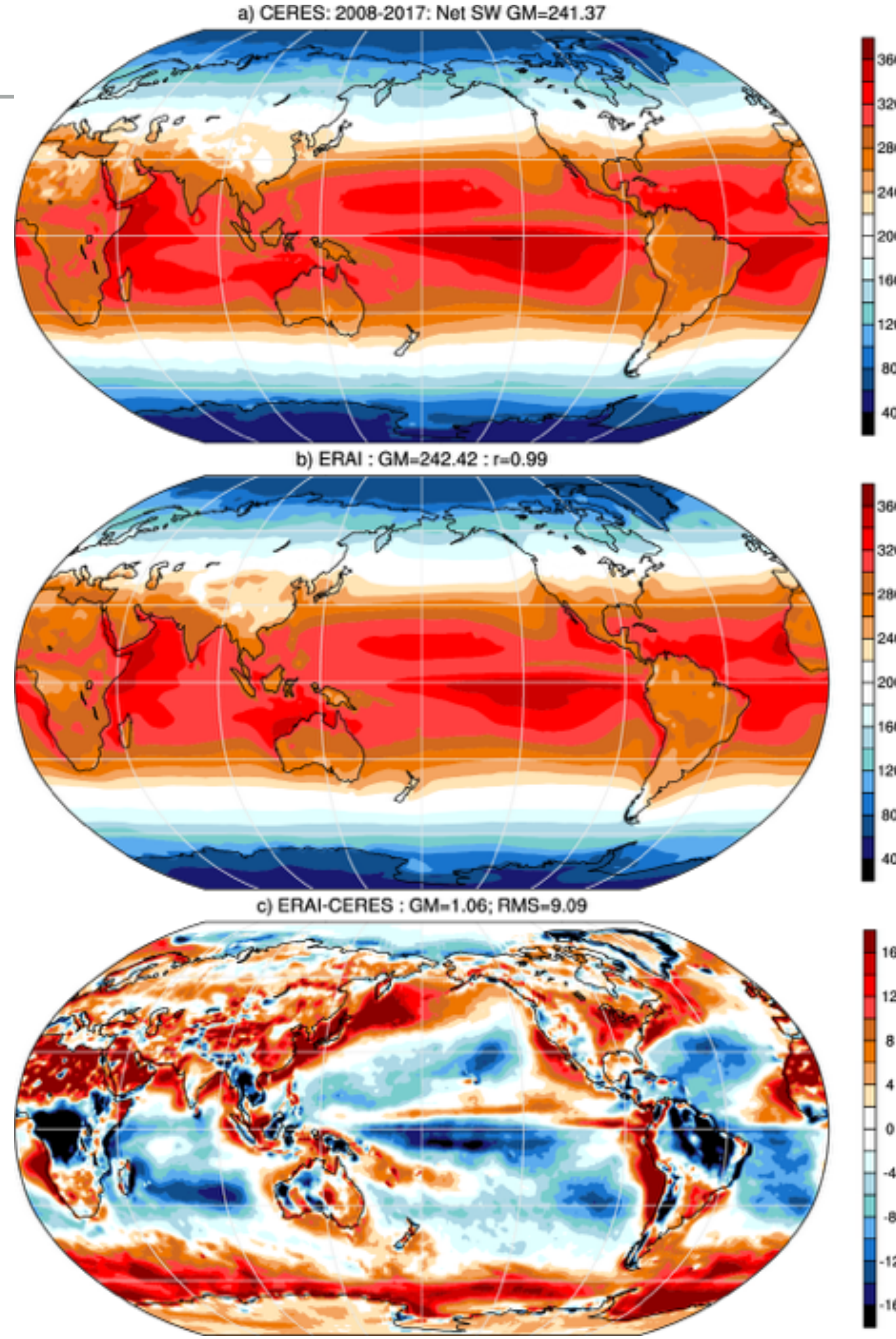
Baselining is accomplished with a 30-yr average (1920-50). Multiple observational datasets are shown. Uncertainty arising from internal variability is estimated from LENS/LME. 60-month smoothing applied.

ADVANCING CMAT: ERA-I

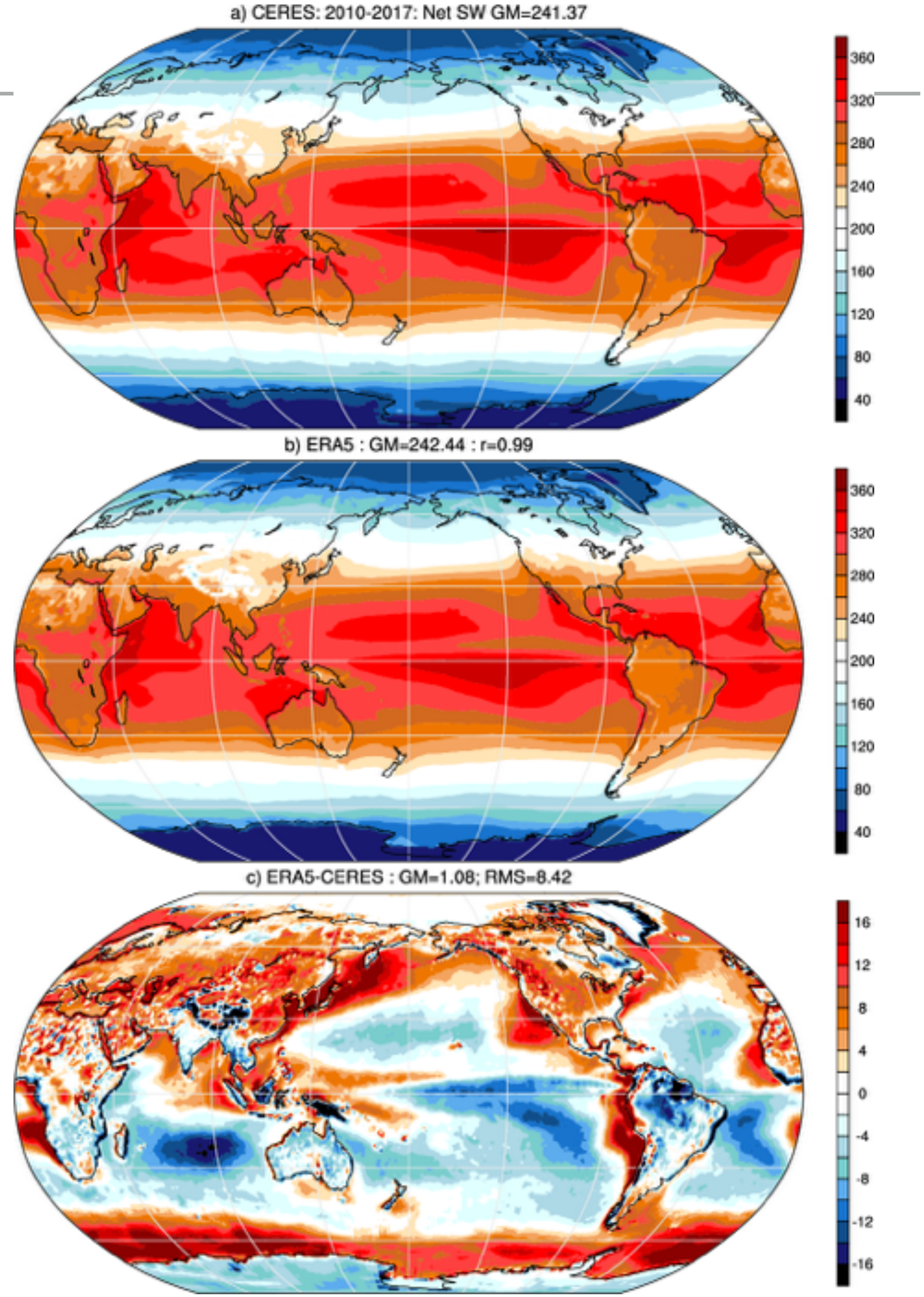
In addition to the advances in CERES discussed at this meeting, other datasets, and particularly reanalyses, continue to improve.

While CMAT uses CERES for TOA fluxes, other fields rely implicitly on ERAI's fidelity in the energy budget and water cycle (E-P,$\nabla \cdot A_E$). The ERA-I release in late 2000s provided a major improvement over earlier reanalyses, but major biases persisted, particularly over tropical land, tropical/subtropical/southern oceans.

Advances in observations, assimilation techniques, and reanalysis models have contributed to a continual improvement of reanalyses.



ADVANCING CMAT: ERA5



Early analysis of ERA5 (2008-present) suggests regional errors in TOA fluxes that are about 50% of those of ERA-I.

Over tropical land, significant additional improvements are evident.

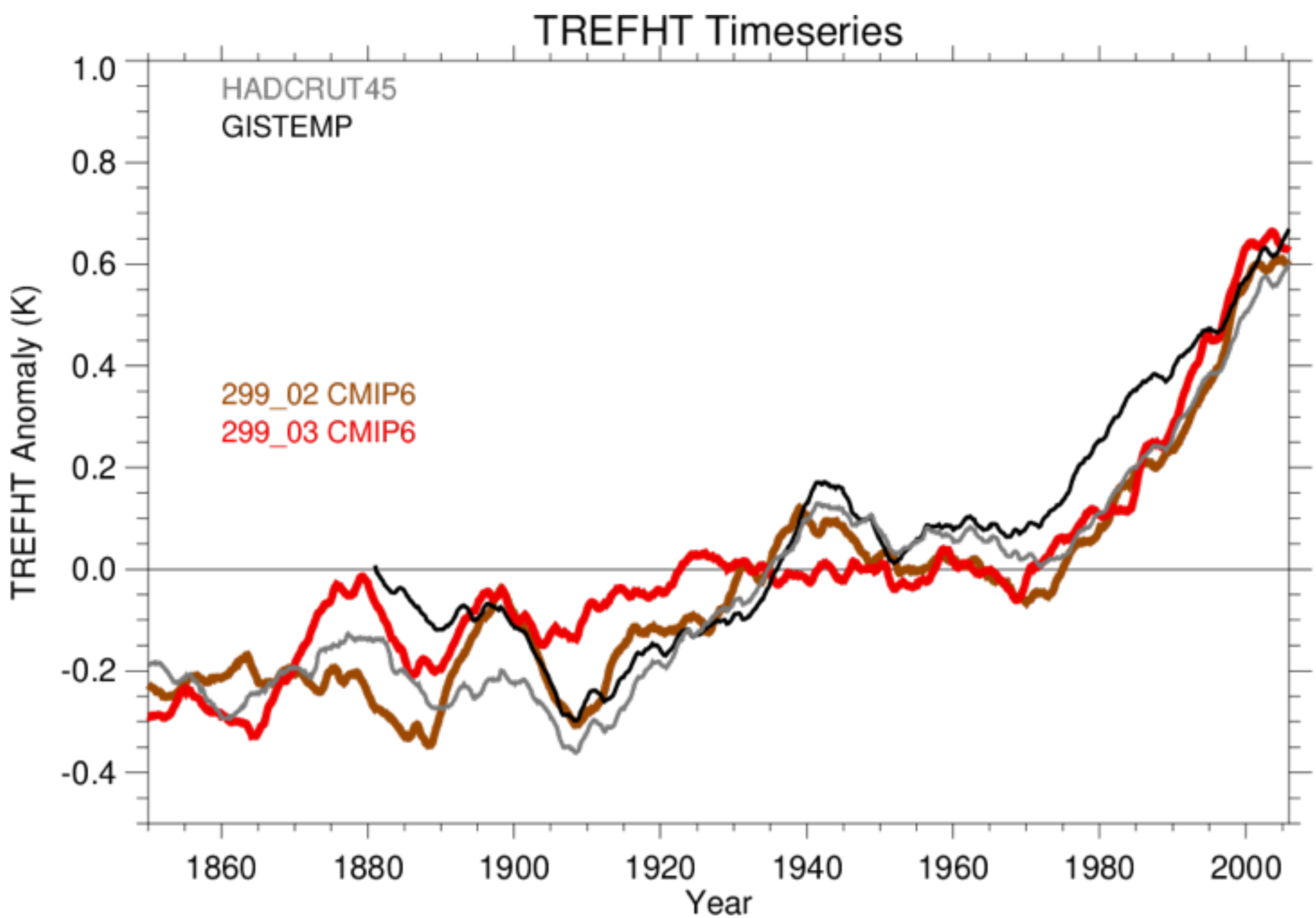
Reductions in error in the reanalysis allow for clearer identification of model bias.

THE CURIOUS CASE OF CESM2

Recent CESM2 simulations have potentially major implications for our understanding of the energy budget.

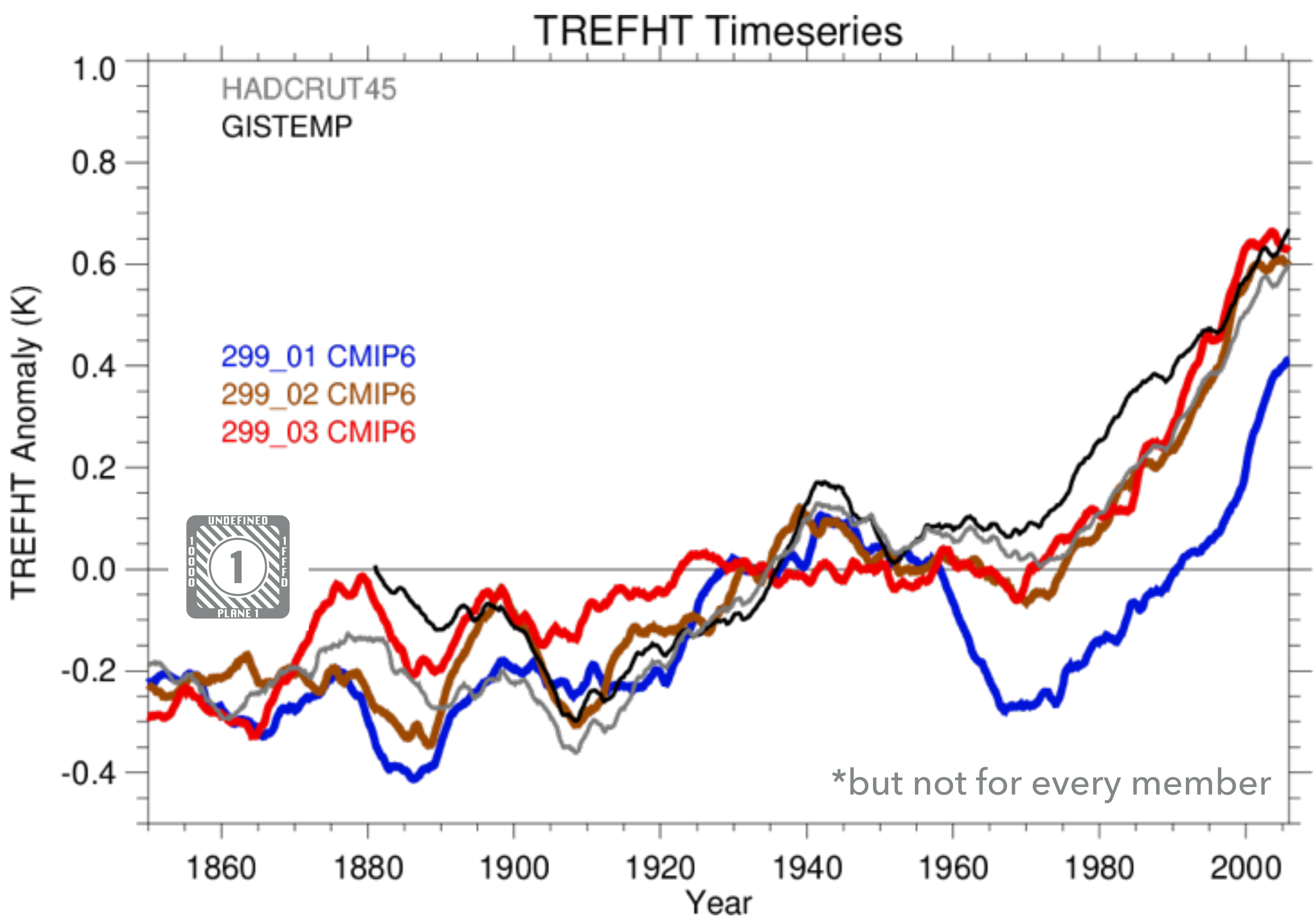
THE CURIOUS CASE OF CESM2

CESM2 matches observed 20th century warming*



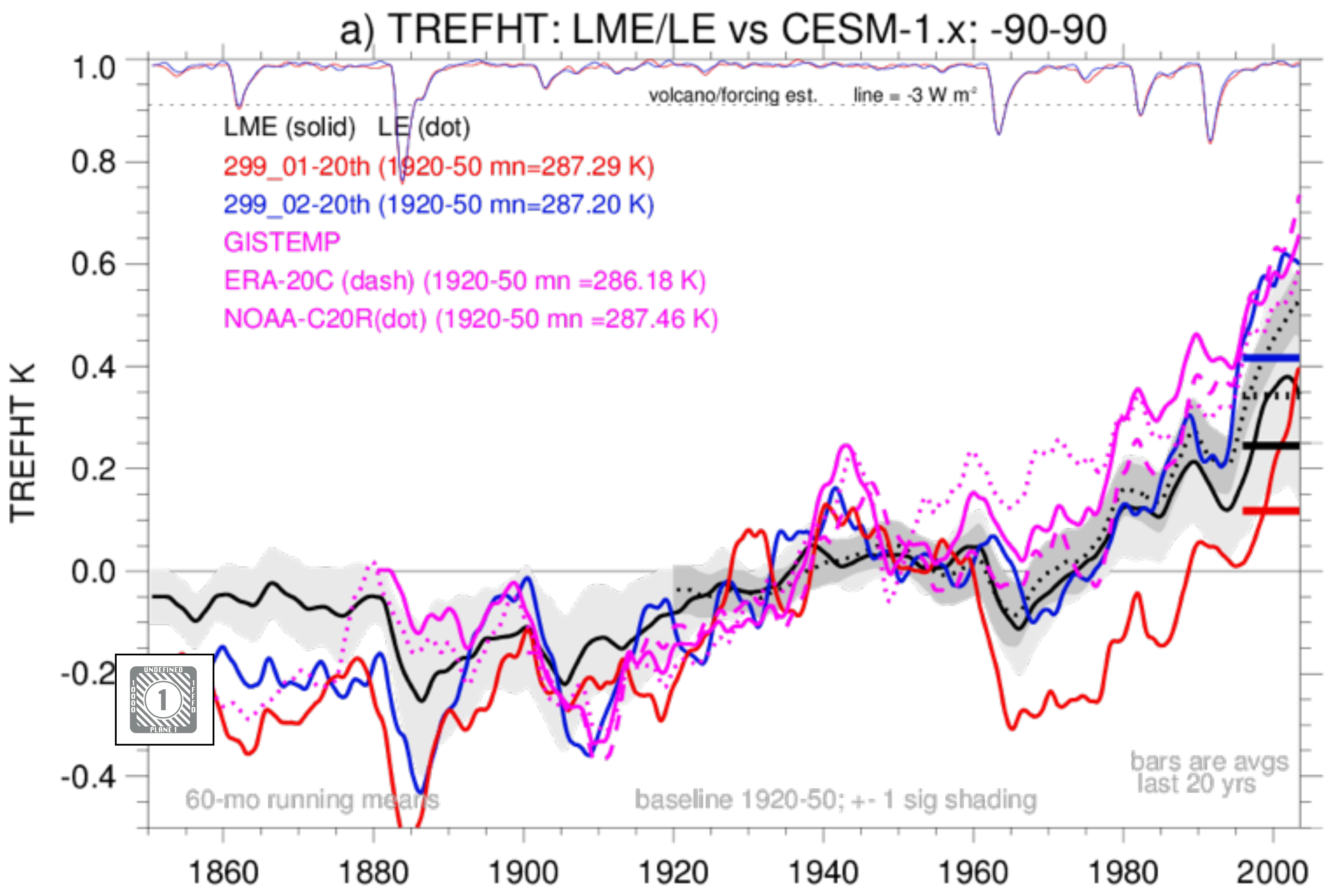
THE CURIOUS CASE OF CESM2

CESM2 matches observed 20th century warming*



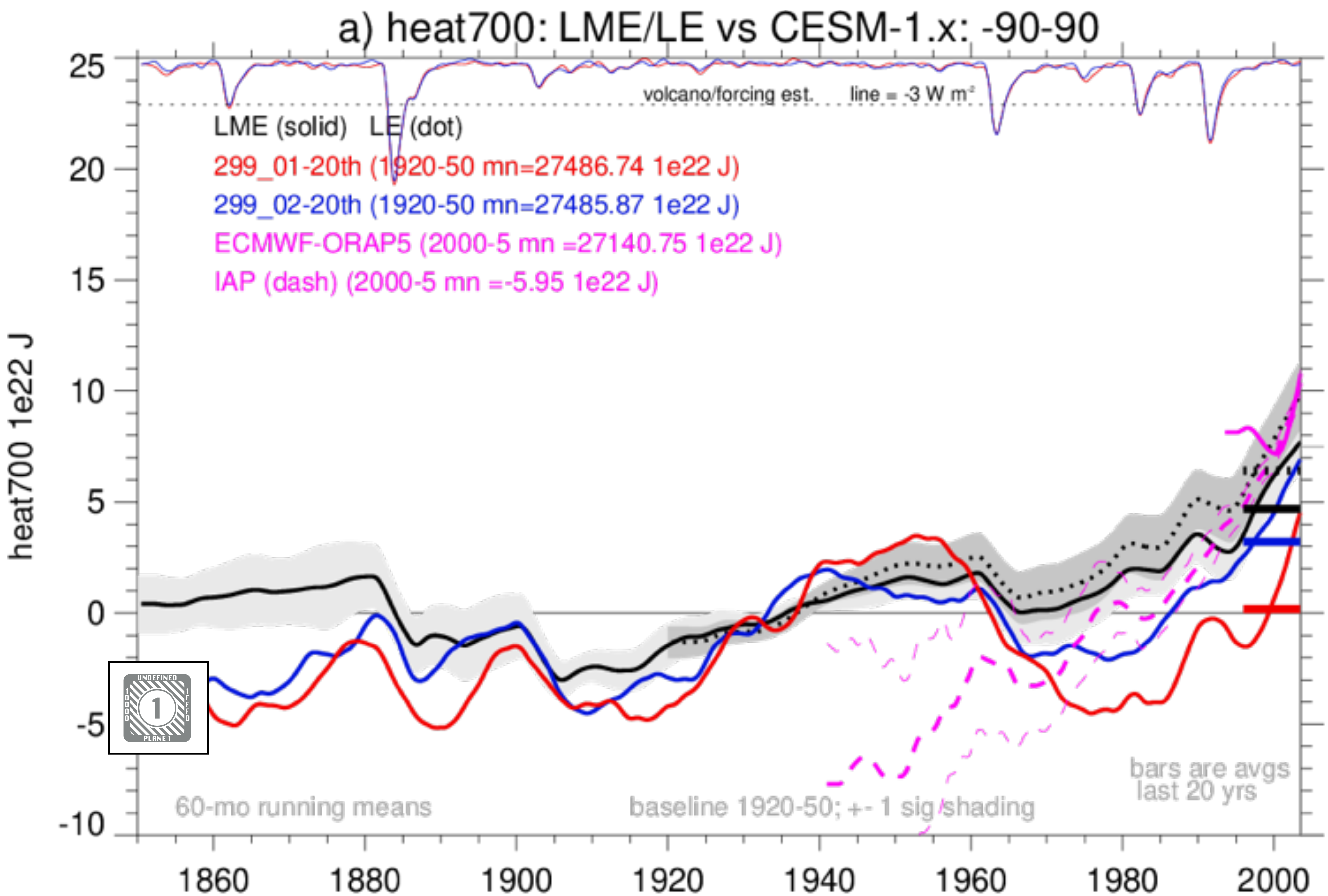
*but not for every member

THE CURIOUS CASE OF CESM2



The internal variability driving the difference is well outside of the bounds of LENS (dark grey shading) despite control runs having variance within ~15%.

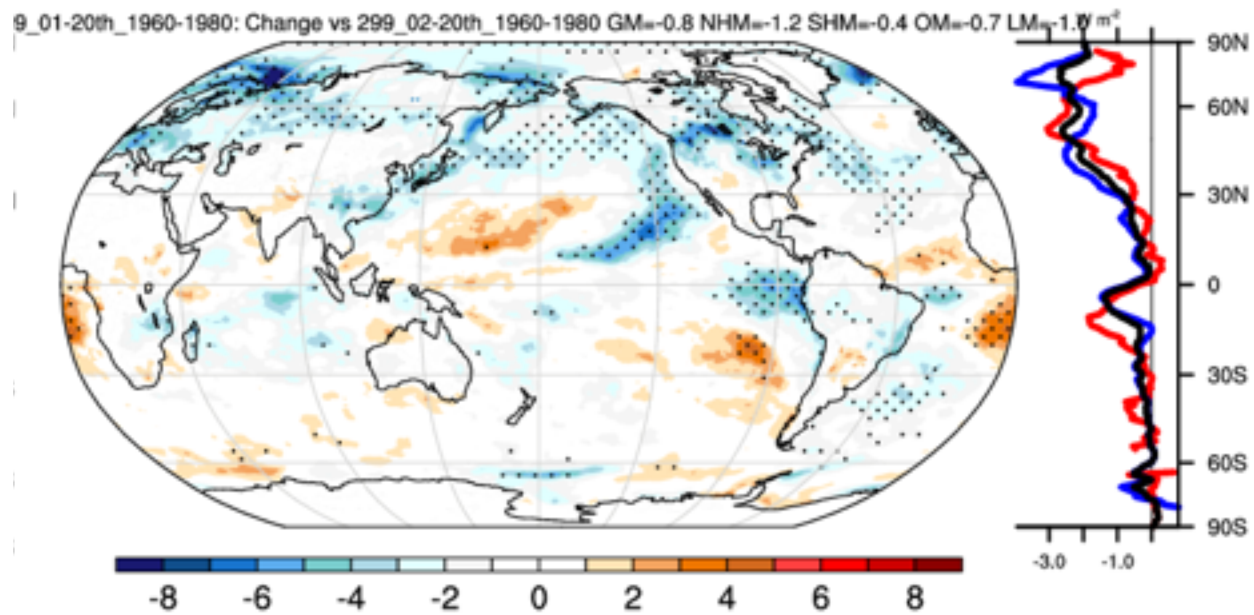
THE CURIOUS CASE OF CESM2



The internal variability has a strong signature in Earth's energy imbalance, unlike that observed (e.g. hiatus). 299_01 cools from 1955-1955! Need to account for both the magnitude and persistence of the cooling.

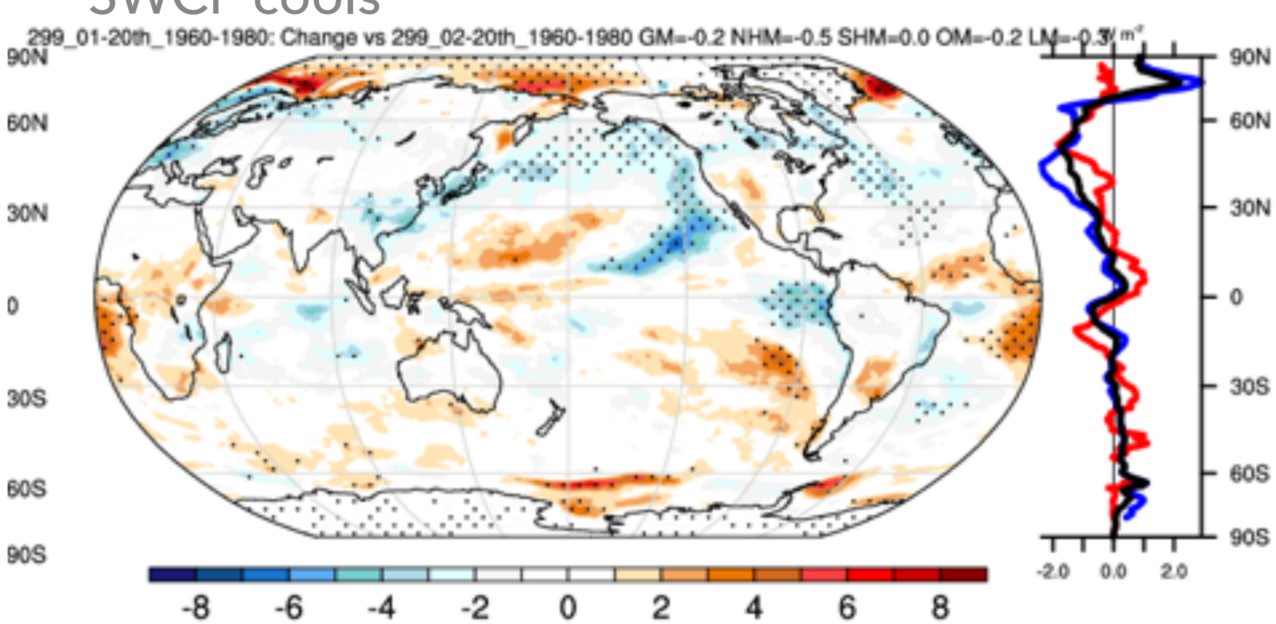
THE CURIOUS CASE OF CESM2

299_01 FSNTOA difference with 299_02

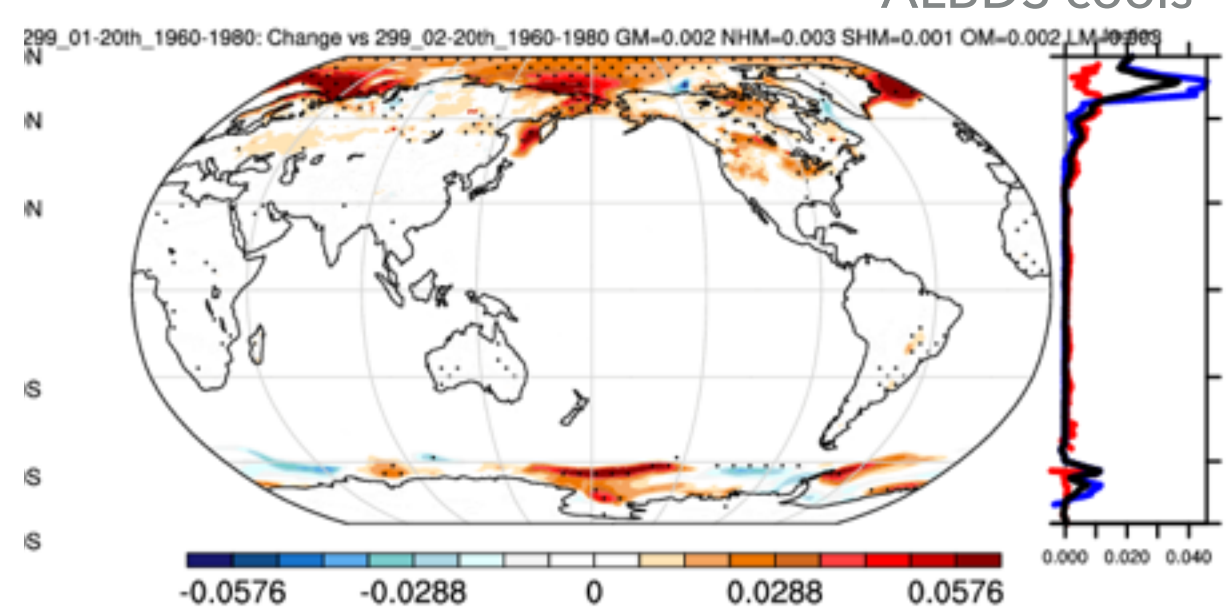


The contrast in Earth's energy imbalance 1960-80 is driven mainly by northern hemisphere albedo contrasts. The run that cools exhibits a substantial increase in albedo over both ocean and land. Why?

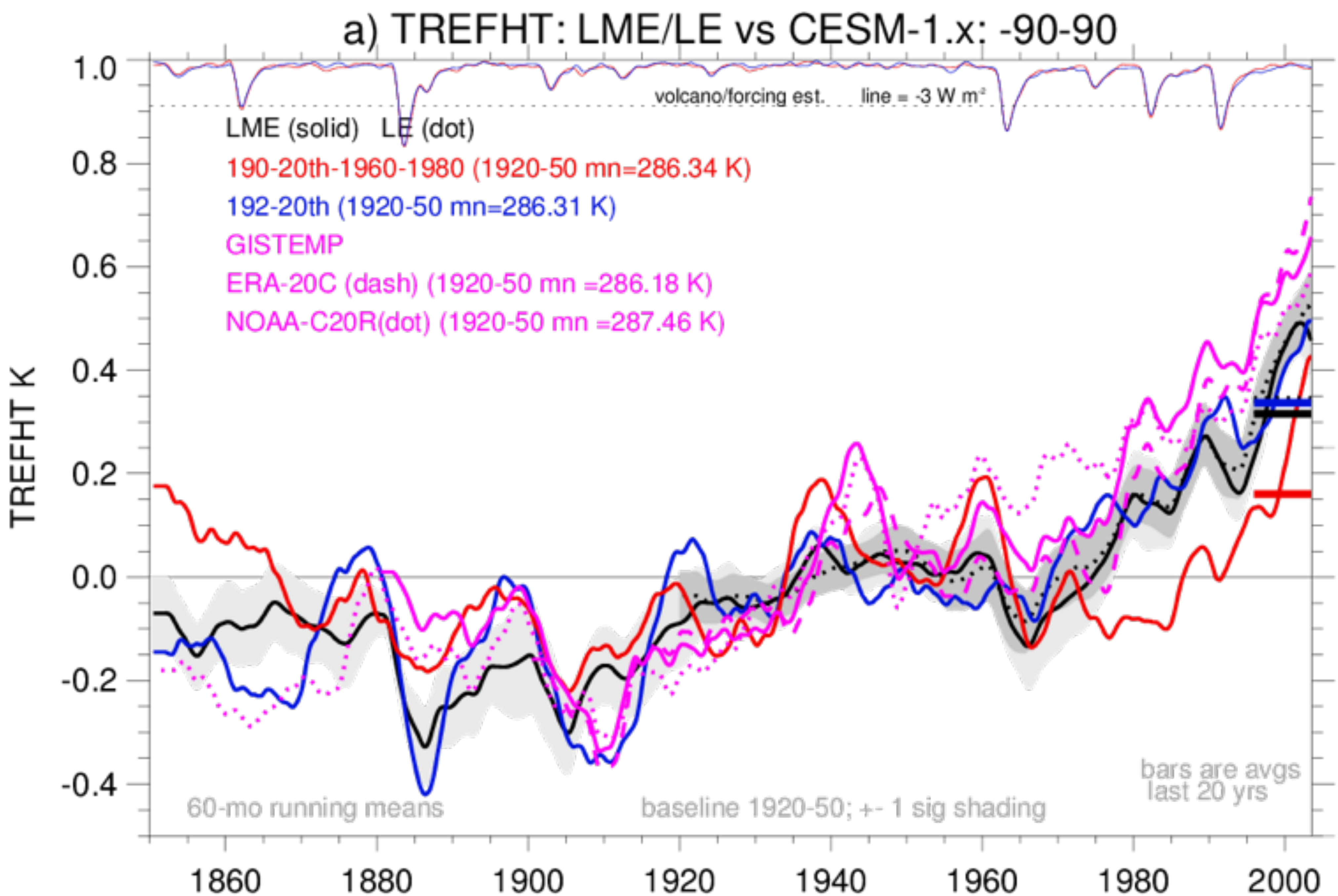
SWCF cools



ALBDS cools

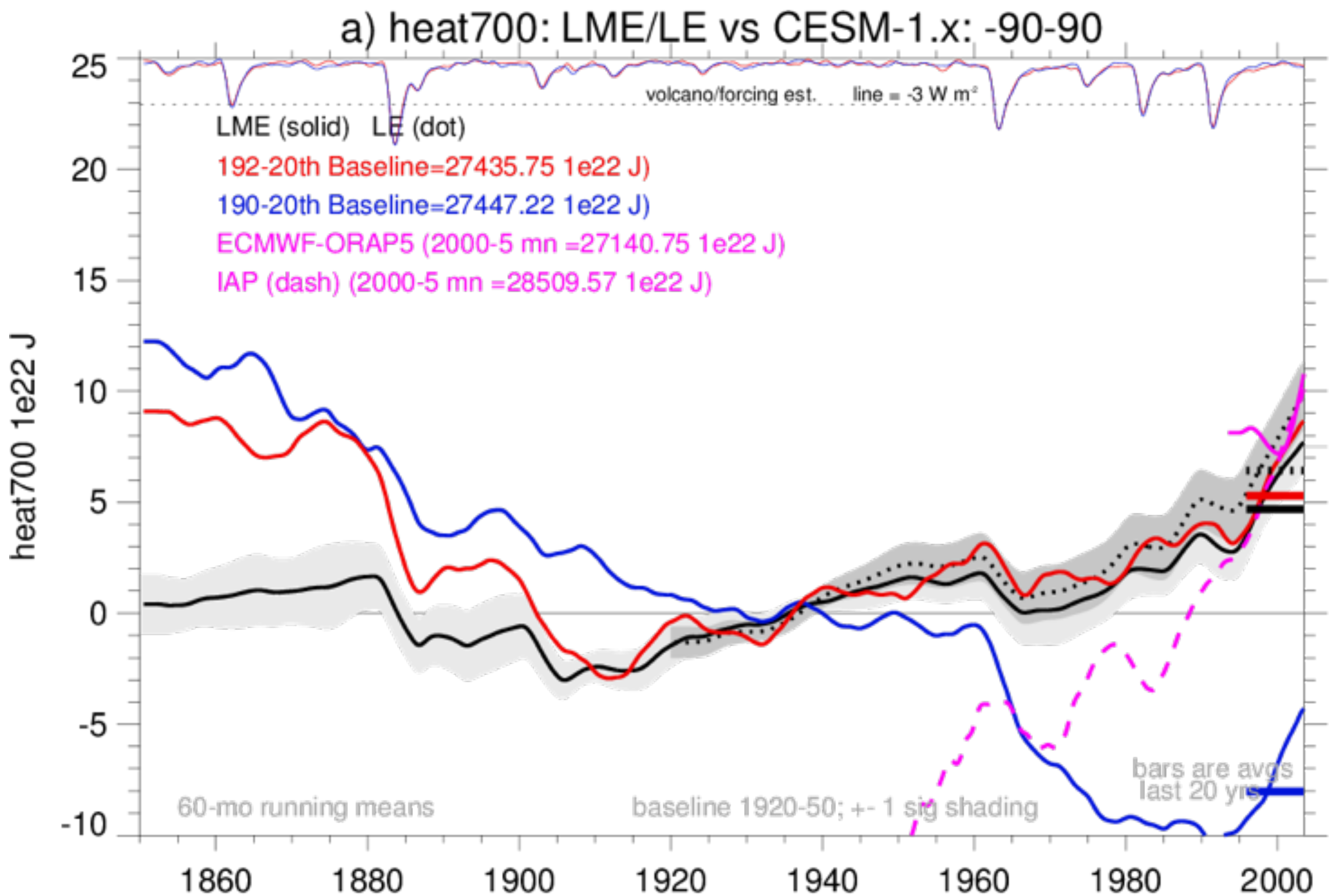


THE CURIOUS CASE OF CESM2



Recall, earlier versions of the model exhibited a similar divergence as do the members of 299.
In the case of 190 vs 192 changes were attributed to the forcing dataset. *Was this justified?*

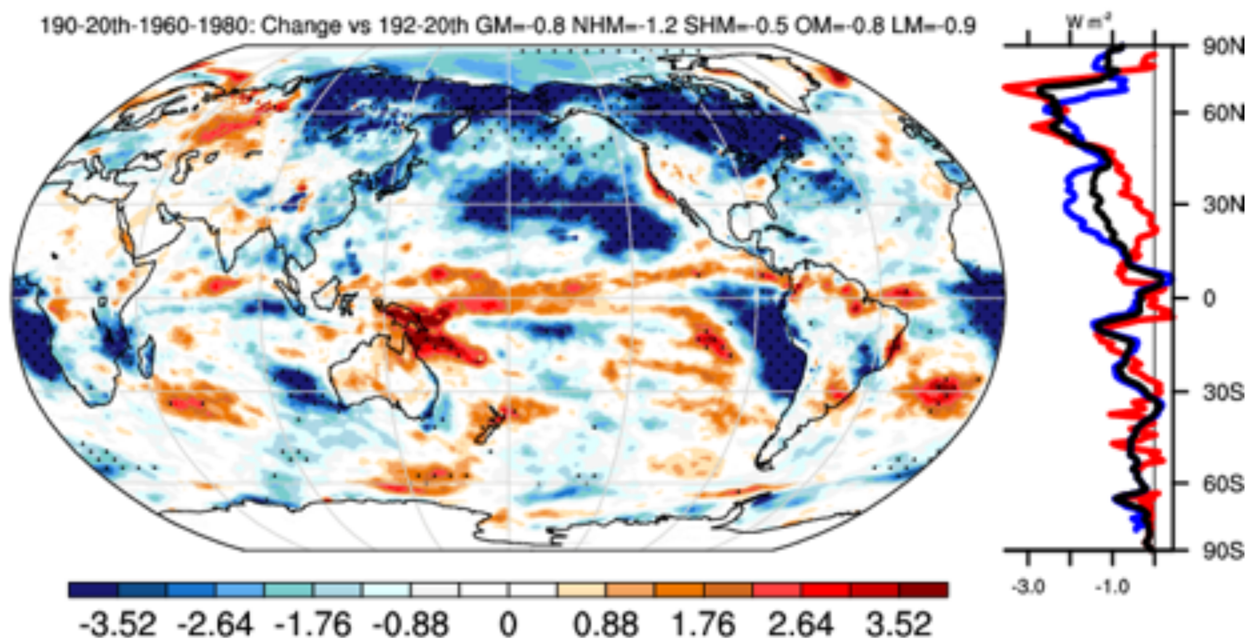
THE CURIOUS CASE OF CESM2



The internal variability has a strong signature in Earth's energy imbalance, unlike that observed (e.g. hiatus). 190 cools substantially from 1955-1995.

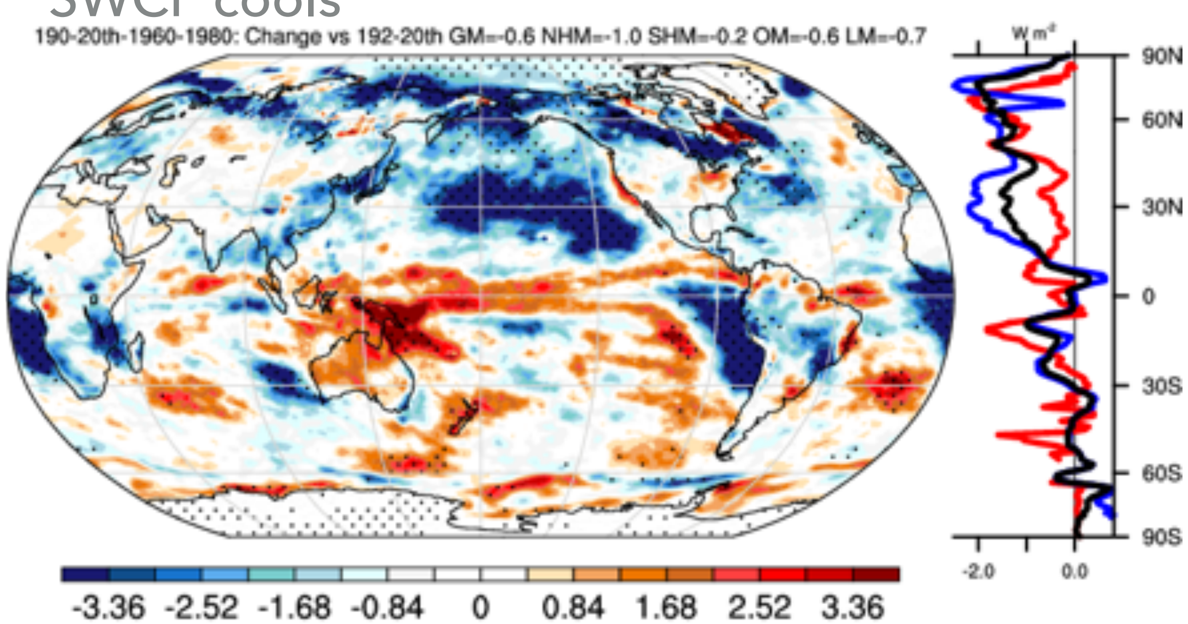
THE CURIOUS CASE OF CESM2

190_01 FSNTOA difference with 192_01

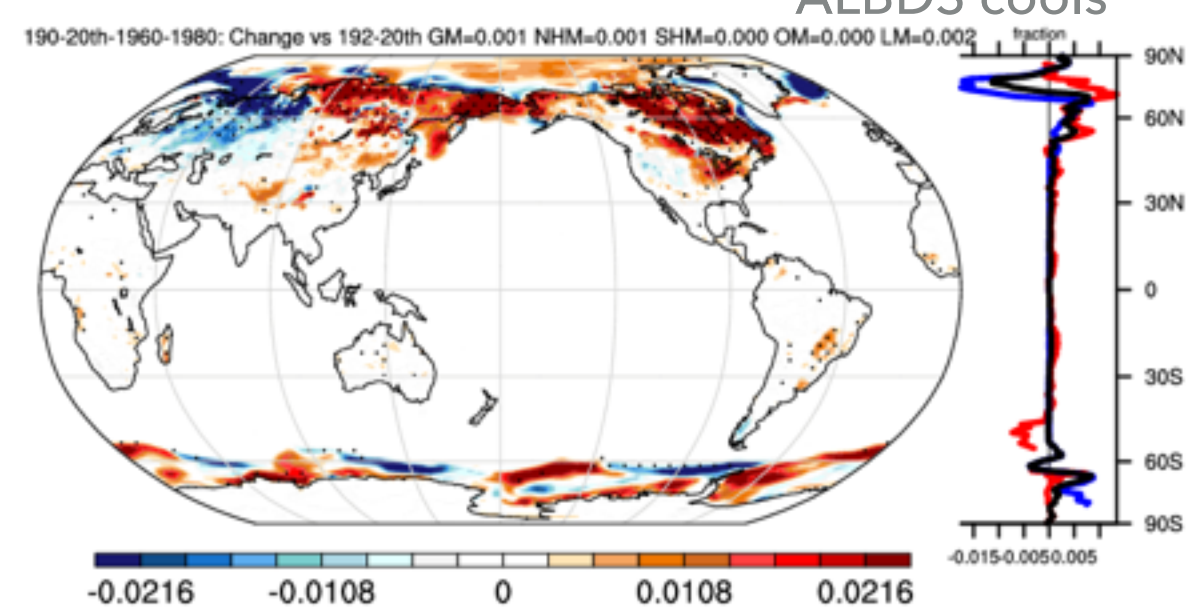


The contrast in Earth's energy imbalance 1960-80 is also driven mainly by northern hemisphere albedo contrasts driven by SWCF and ALBDS.

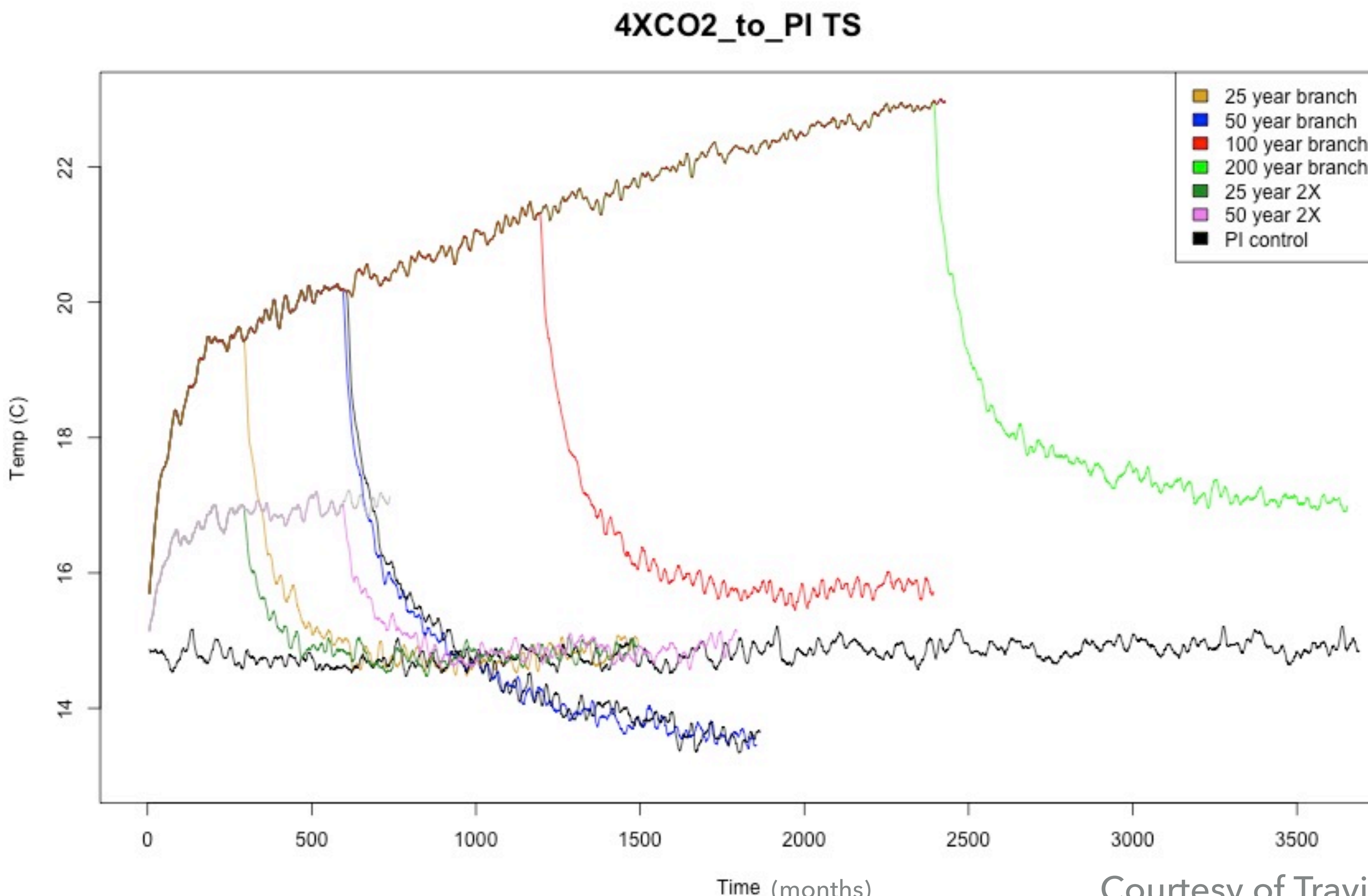
SWCF cools



ALBDS cools



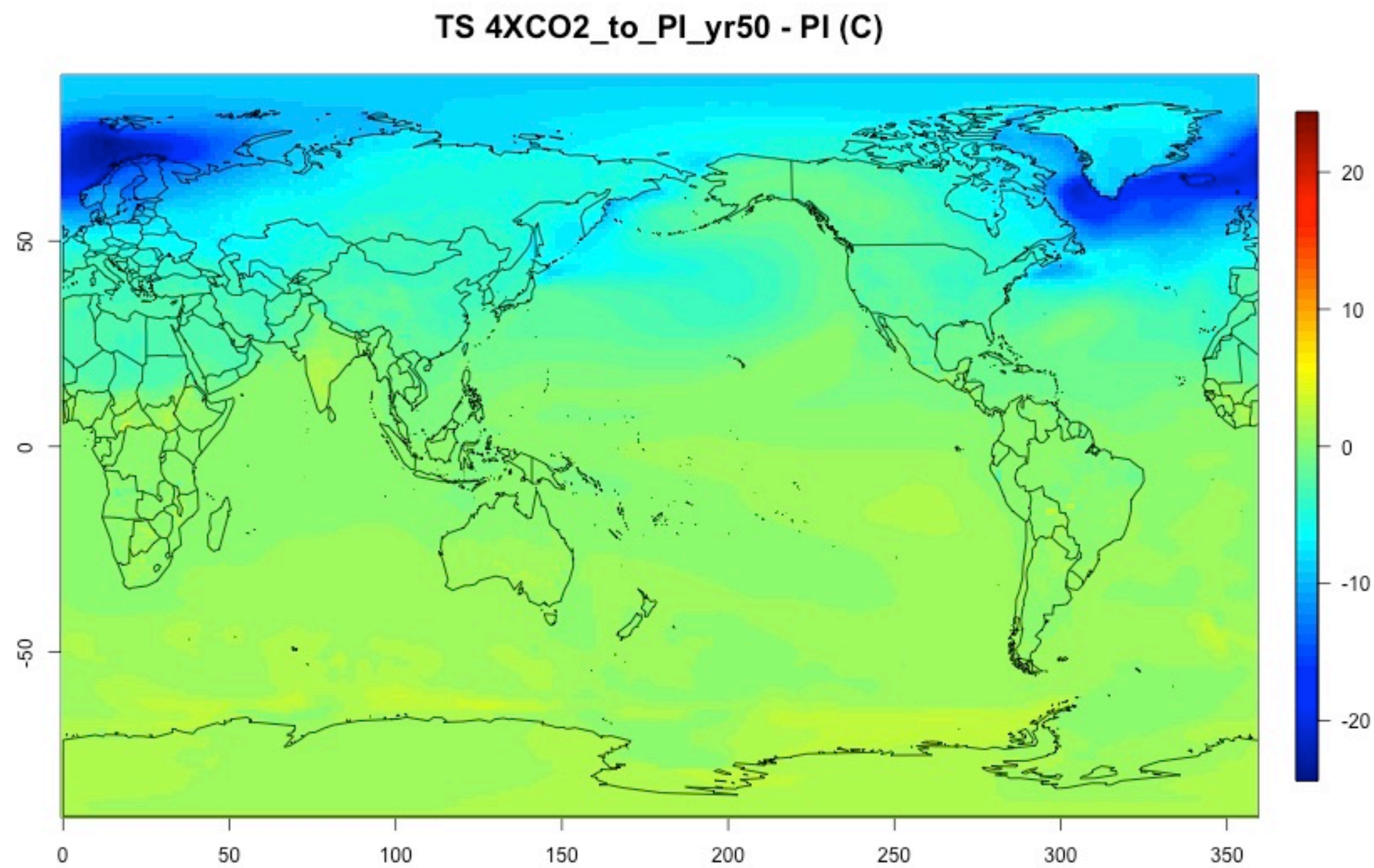
THE CURIOUS CASE OF CESM2



Courtesy of Travis Aerenson

4xCO2 runs that return to 1xCO2 exhibit hysteresis and apparently multiple stable states.

THE CURIOUS CASE OF CESM2



Courtesy of Travis Aerenson

CONCLUSIONS

- CMAT has been developed to provide a comprehensive and objective model evaluation tool based on multiple physically relevant climate metrics involving both the climatological mean and aspects of variability (seasons/ENSO).

- A focus is given to fields involving the energy budget due to its strong physical ties to transient climate variability and change. It's components (e.g. OHC) are also more robust to internal variability than are trends in surface temperature.

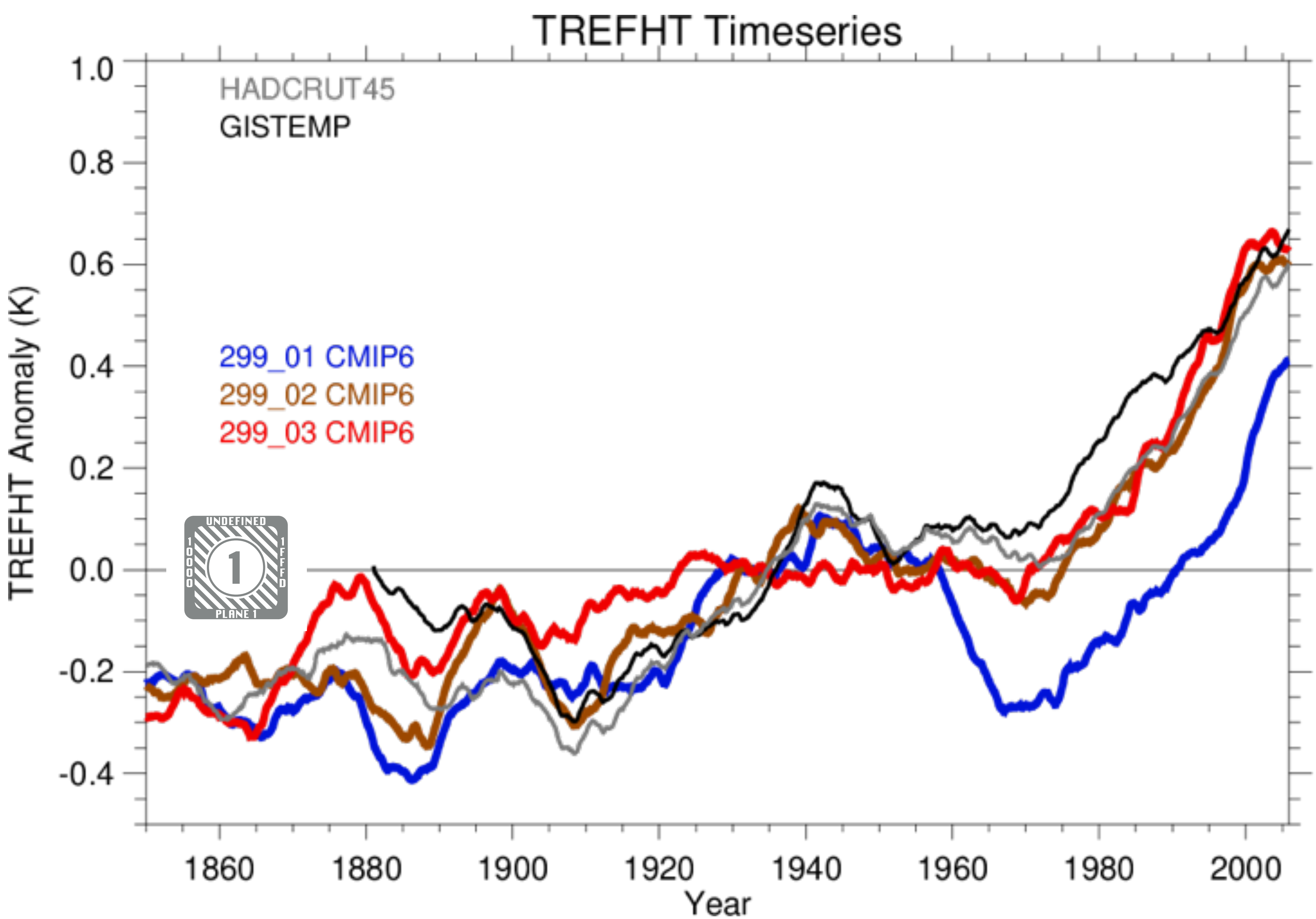
- CESM2 scores the highest in CMAT of all climate models tested. Its cloud scheme corrects a long-standing systematic bias in the latitudinal structure of SW cloud forcing present in other CMIP3 and CMIP5 models.

- Yet CESM2 exhibits some astonishingly unique behavior with apparent strong dependencies of EEI on internal variability. Some 20th century transient ensemble members exhibit a persistent negative Earth energy imbalance in the later half of the 20th century due to increases in albedo at high latitudes. Others do not.

- The model exhibits nonlinearities, such as hysteresis in idealized 4xCO2 perturbation experiments. What is the physical basis for these?

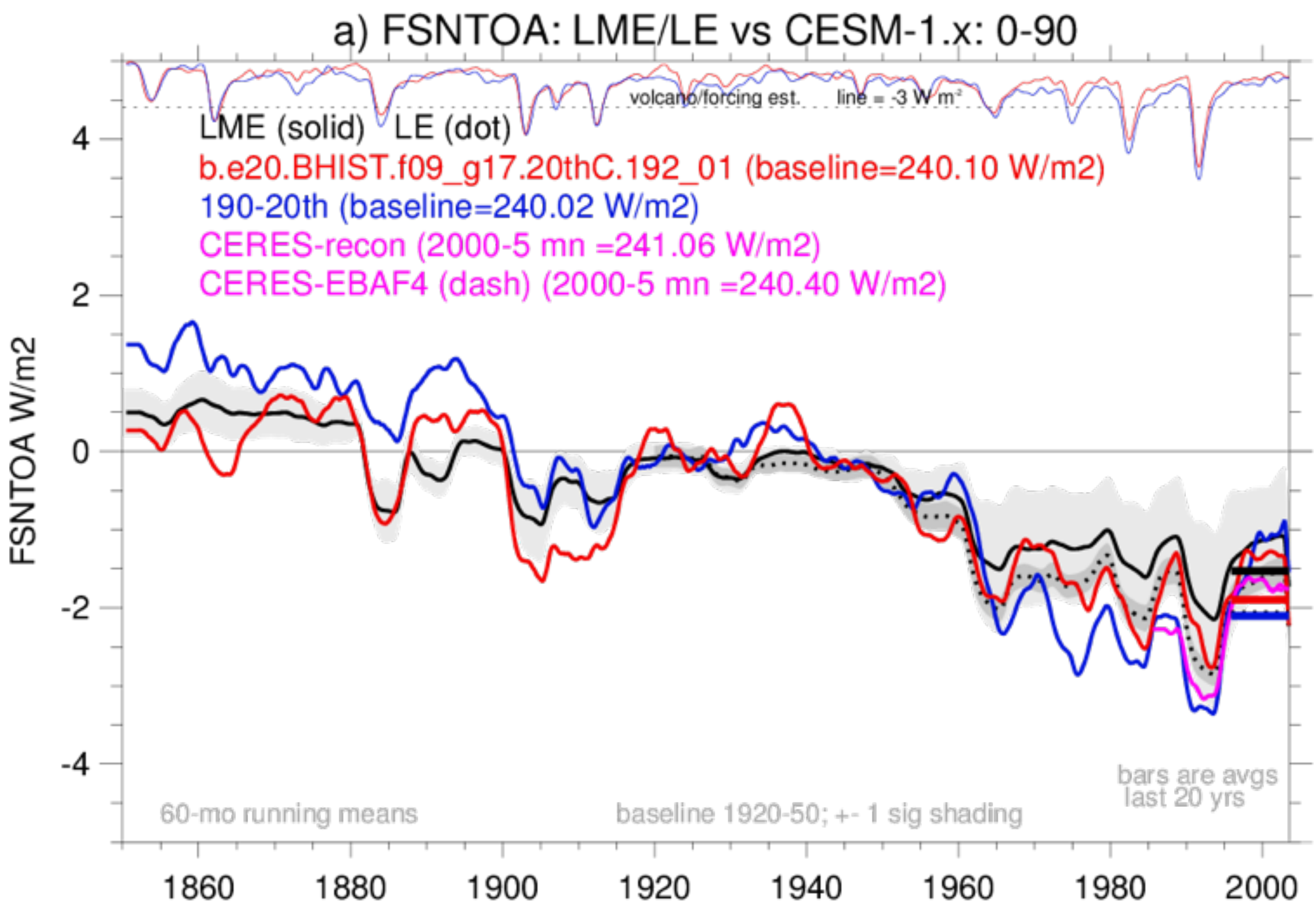
WHAT IF CESM2 IS RIGHT?

Did late 20th century climate warming depend strongly on internal processes? If so, what would the broader implications have been for climate science … and policy?

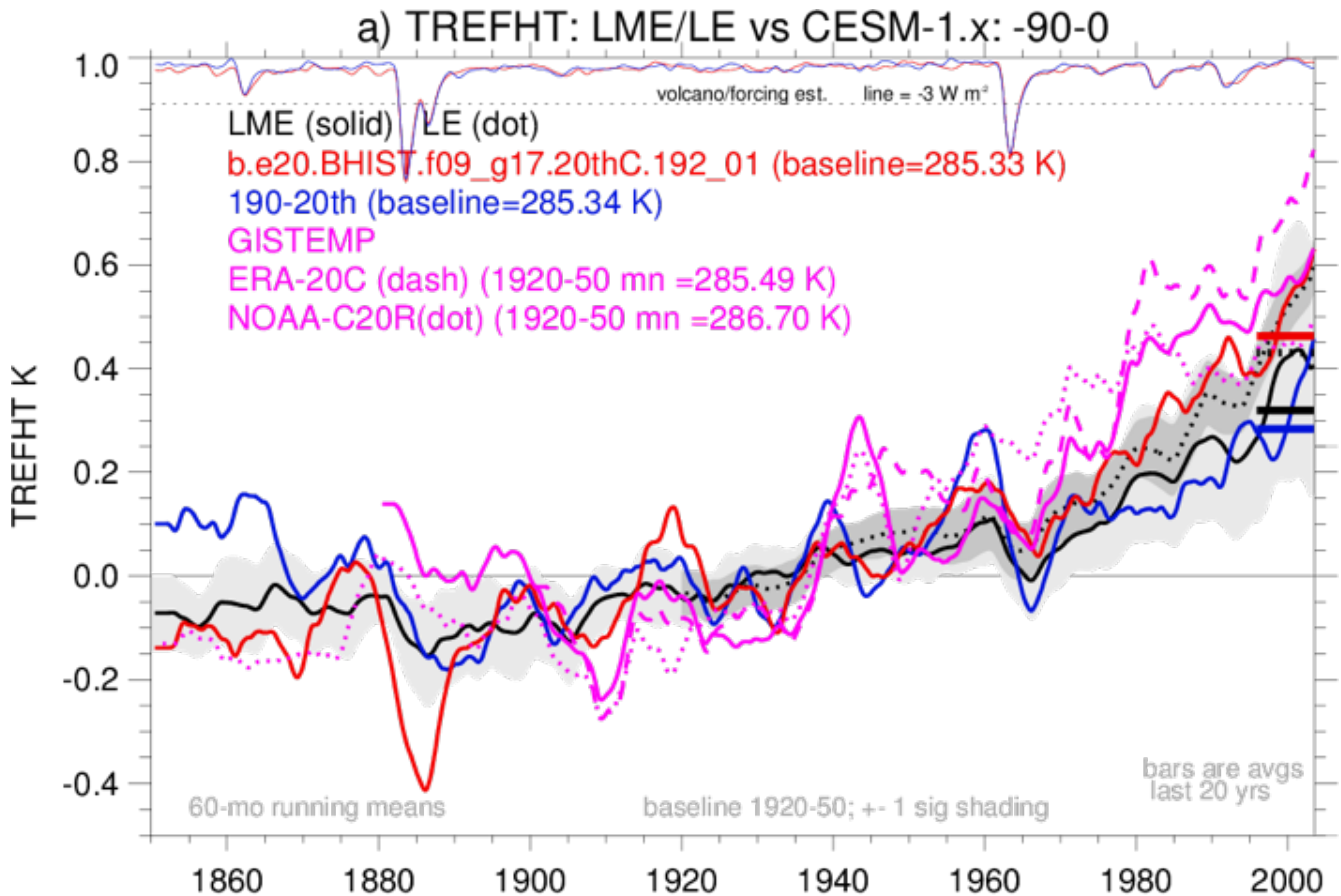


END

CMAT SUMMARY PAGES: TIMESERIES

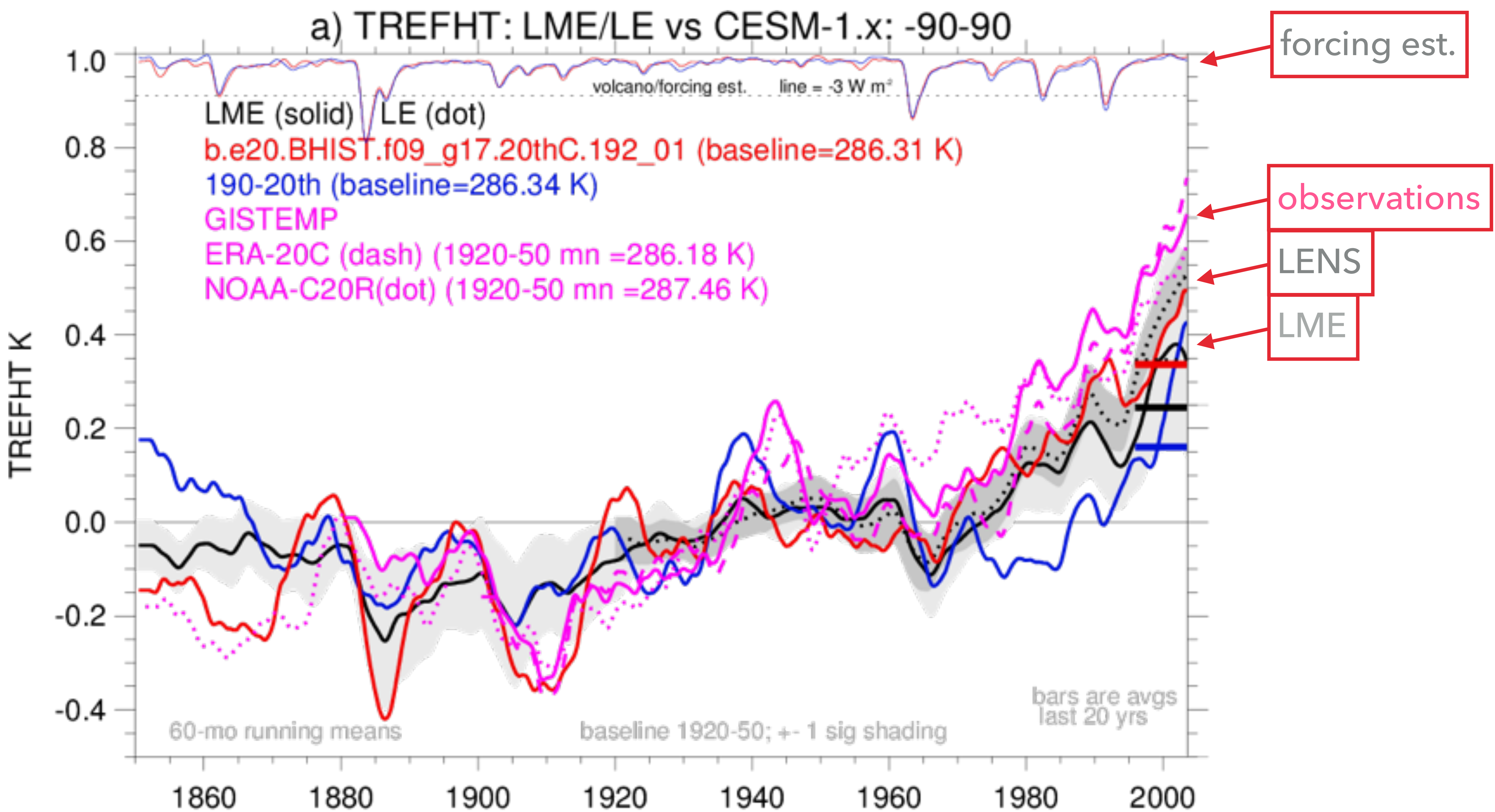


CMAT SUMMARY PAGES: TIMESERIES



- SH warming in **192** is on par with LENS, in **190** is on par with LME

CMAT SUMMARY PAGES: TIMESERIES



• LENS often used as a benchmark but warming in LENS is likely a bit too small

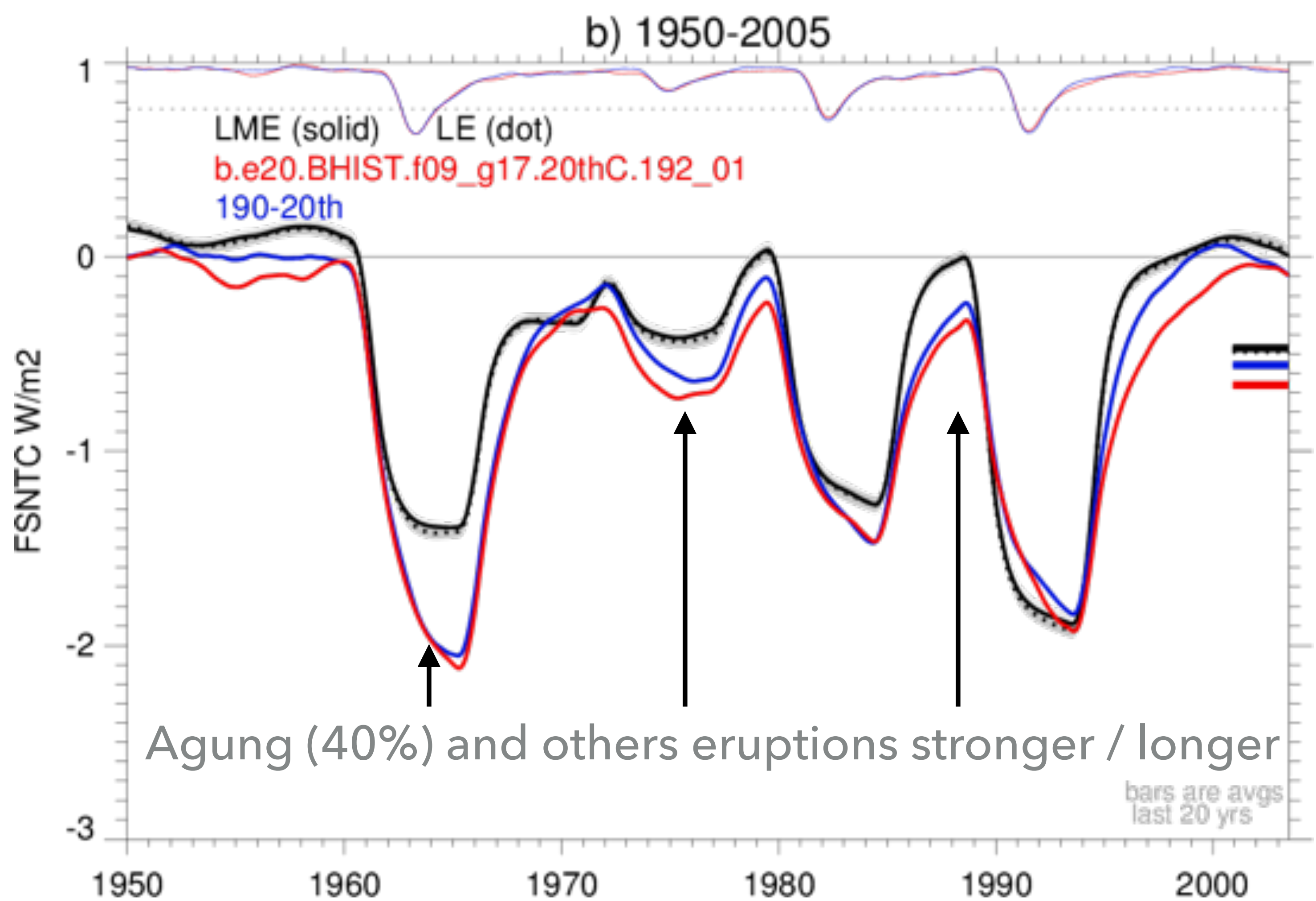
2 key science questions: 1) why does LENS warm more than 190, 2) why does 192 warm more than 190

• Need to explain not only why it fails to warm 1965-95 but also why it warms so rapid in 1990s

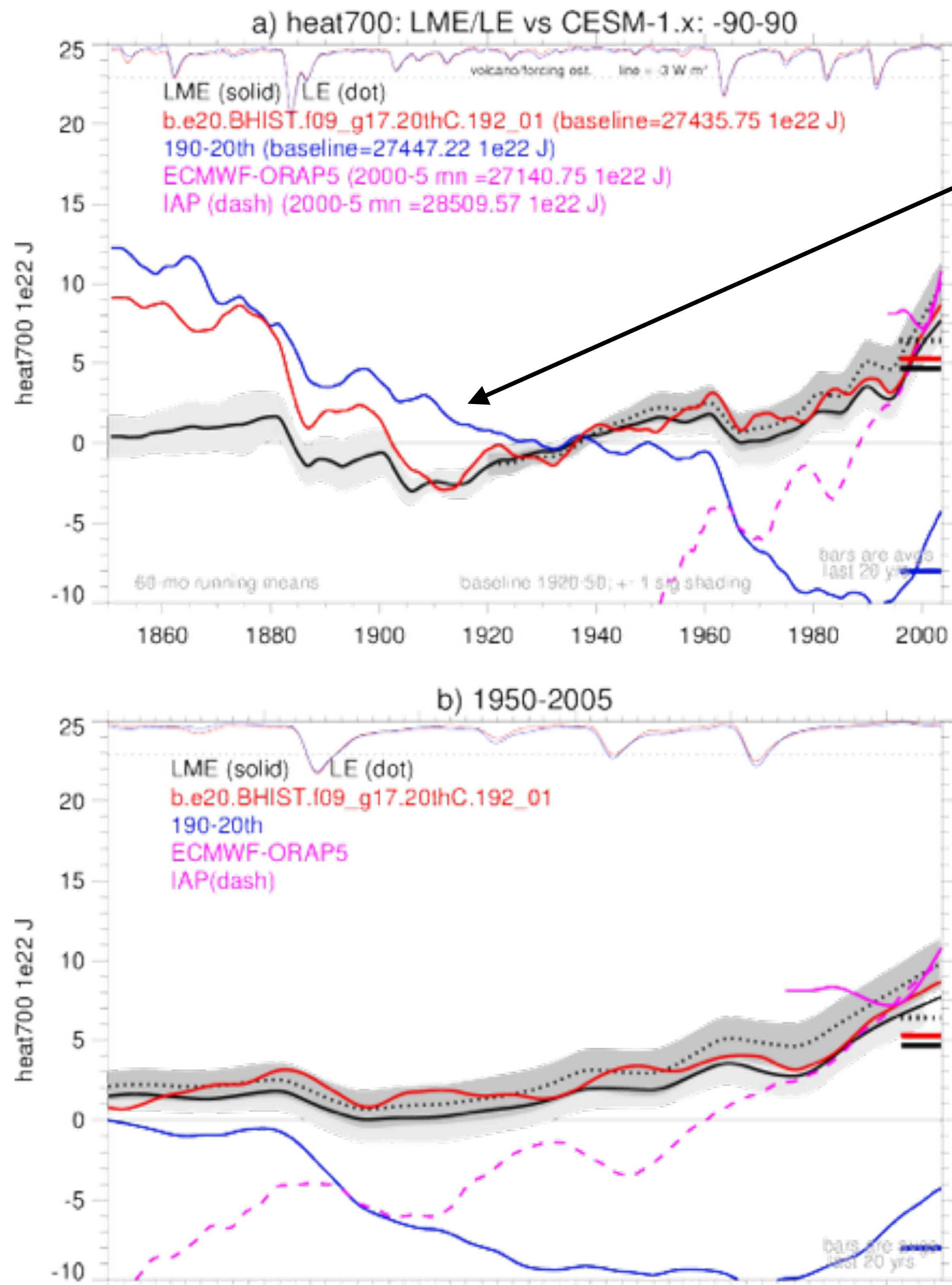
• Baseline temperature is almost identical in 190/192, as is warming by 2005.

190_01 LACK OF WARMING VS LENS : INFLUENCE#1 : ROLE OF VOLCANIC FORCING

Clear Sky Net TOM Solar Flux: Ocean 30N-S



190_01 LACK OF WARMING : INFLUENCE #2 : OCEAN DRIFT



Long term cooling trend in OHC 0-700m is indicative of heat being lost from the upper ocean and stored in the deep ocean.
(**0.15 Wm^{-2}**=50% of late 20th C in obs)

Relevant mainly to trend. Perhaps a secondary influence on 1960-2005. Does not seem to explain why 190 and 192 diverge.

190_01 LACK OF WARMING : INFLUENCE #3 : SNOW COVER FEEDBACK

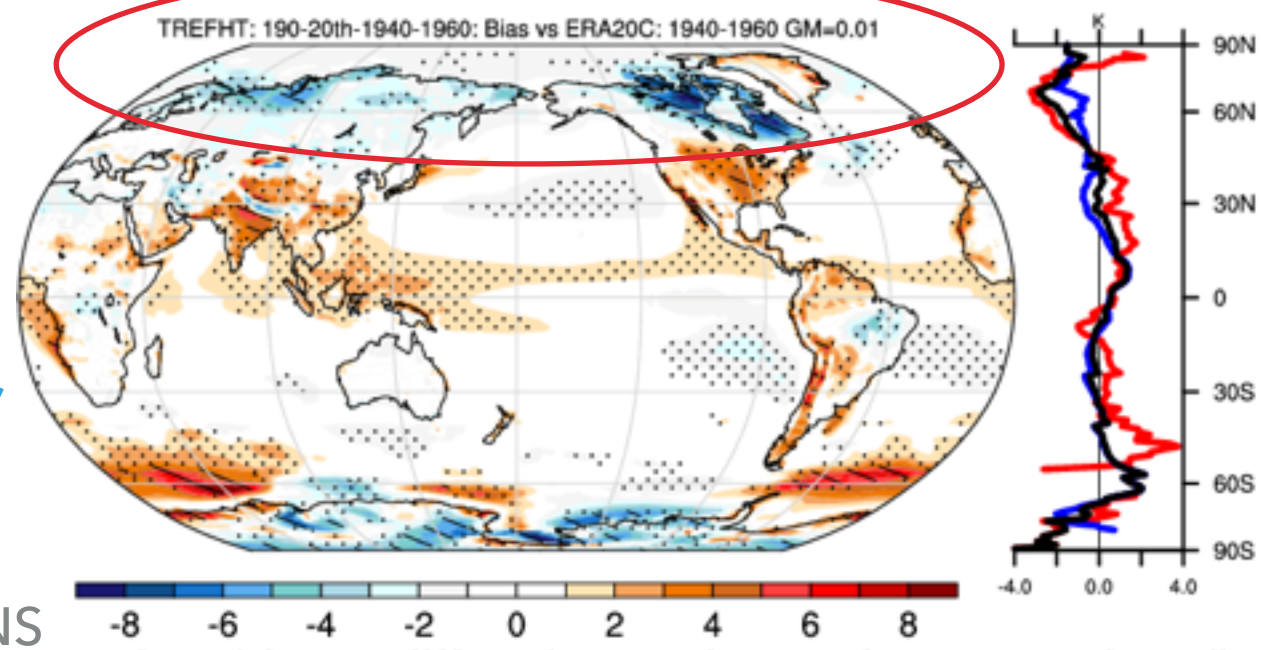
1940-60: BASE STATE IS BIASED COLD (4-6K) IN REGIONS OF KEY ALBEDO FEEDBACKS DURING LACK OF WARMING (ERA20C used for obs)

VOLCANIC FORCING+ EMISISONS +AIE INDUCE A COOLING PULSE

COOLING EFFECT OF AEROSOLS AND COLD MEAN STATE ALLOWS SNOW COVER TO PERSIST INCREASINGLY THROUGH SUMMER FROM 1960-95; CLOUD AMOUNT INCREASES; AT TIME OF PEAK ANNUAL SOLAR FLUX

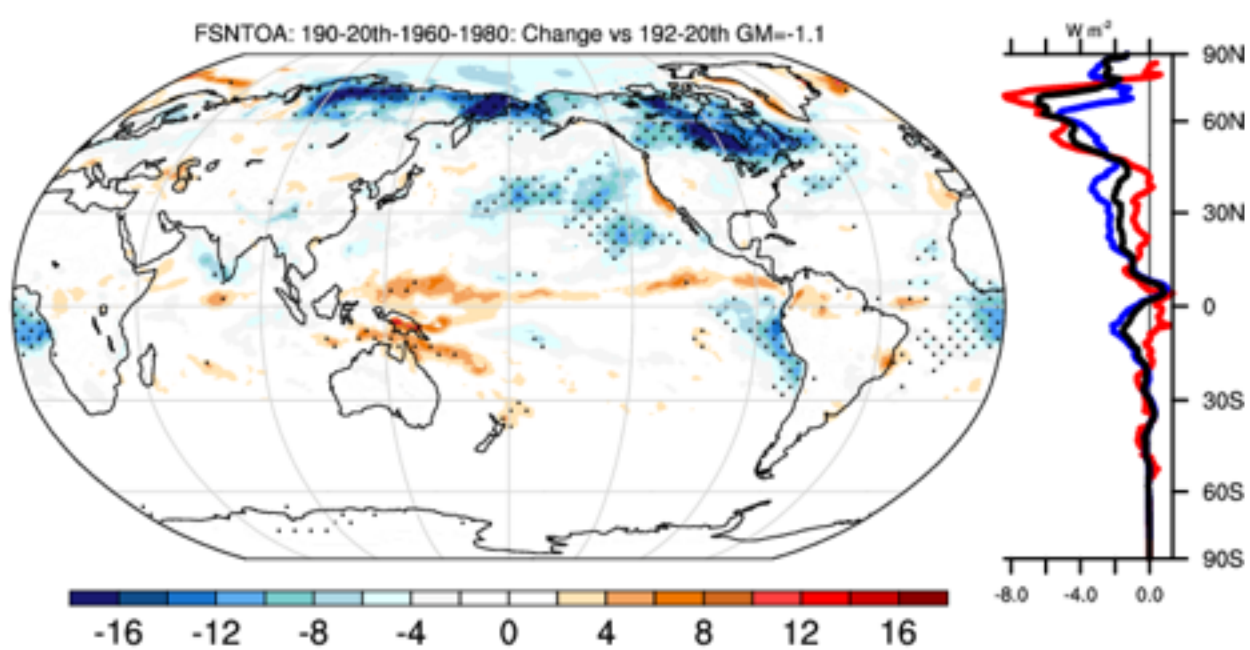
∴MODEL FAILS TO WARM

TREFHT: 190 VS ERA20-C: JJA 1940-60



hatching = diffs > internal var+obs uncertainty (lef

FSNTOA: 190 VS 192: JJA 1960-80



t column only); stippling = diffs > internal var only

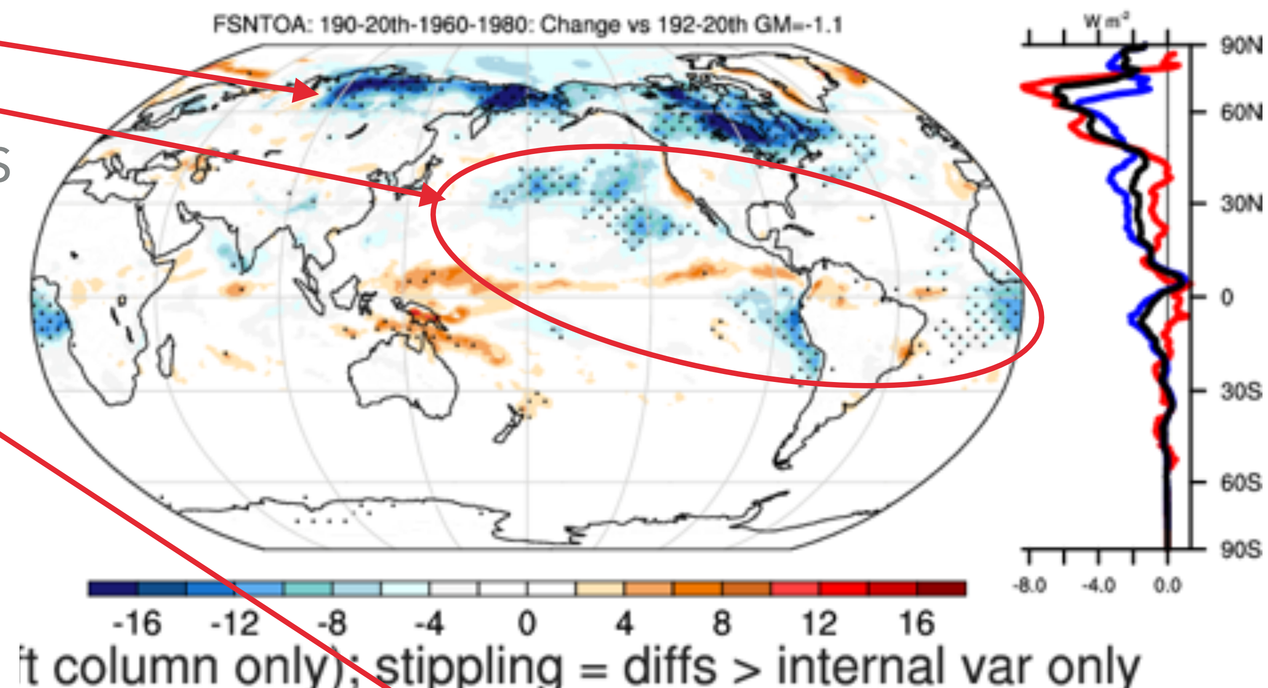
190_01 LACK OF WARMING : INFLUENCE #3 : SNOW COVER FEEDBACK

FROM 60-80, BOTH FORCING AND FEEDBACK EFFECTS ARE CLEAR.

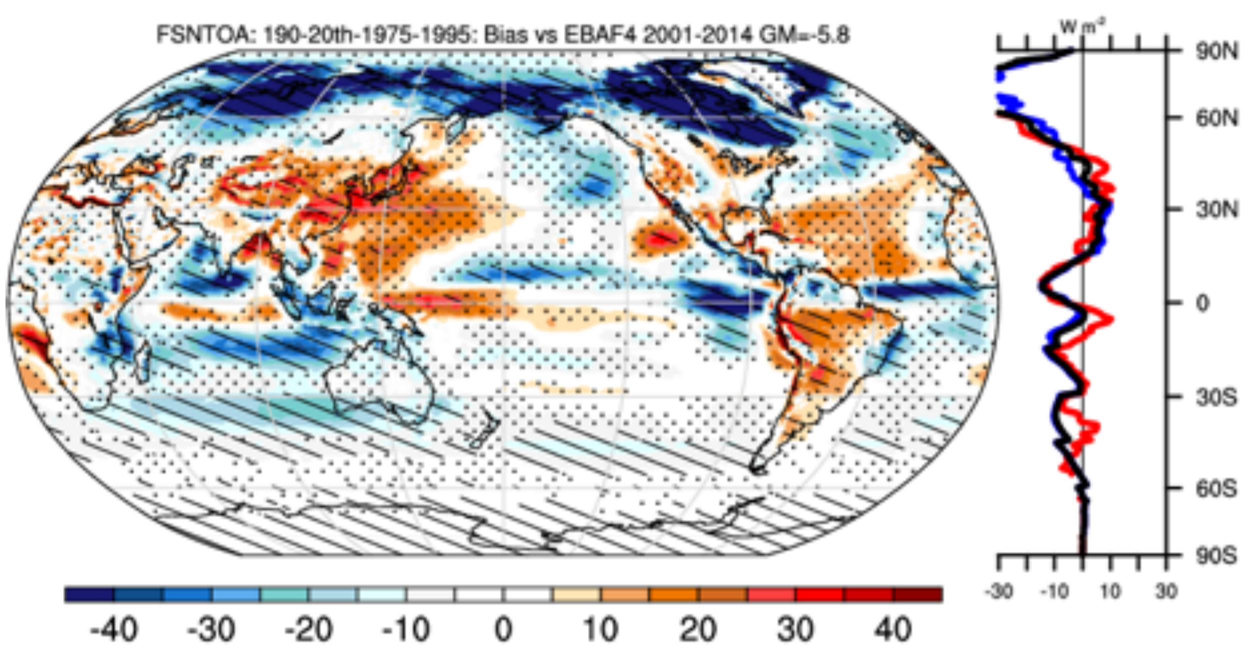
FROM 75-95, ALBEDO BIASES AT HIGH LATITUDES DOMINATE THE GLOBAL DIFFERENCE BETWEEN 190/192.

FLUXES ALSO DEPART SIGNIFICANTLY FROM CERES OBSERVATIONS IN BOTH 190 AND 192. ∴**THIS IS LIKELY A MAJOR MODEL BIAS ENABLED BY A COLD ARCTIC, NOT MERELY A FEEDBACK TRIGGERED BY EXCESSIVE FORCING.**

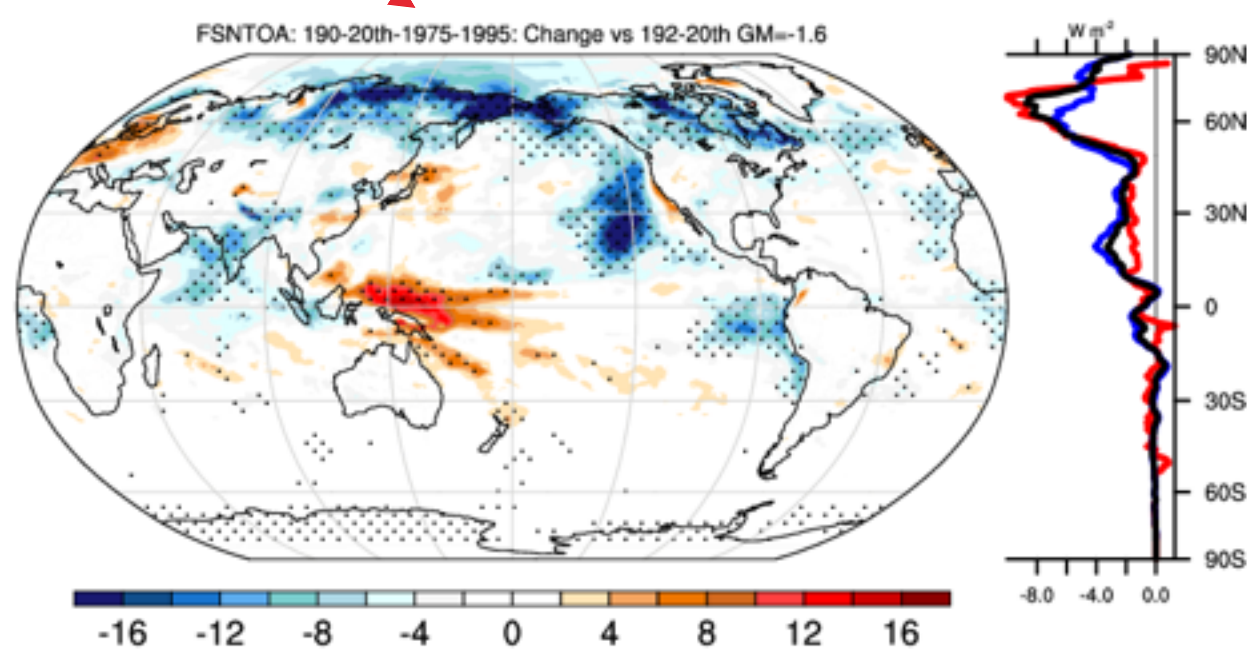
FSNTOA: 190 VS 192: JJA 1960-80



FSNTOA: 190 VS EBAF4(2000-14): JJA 1975-95



FSNTOA: 190 VS 192: JJA 1975-95

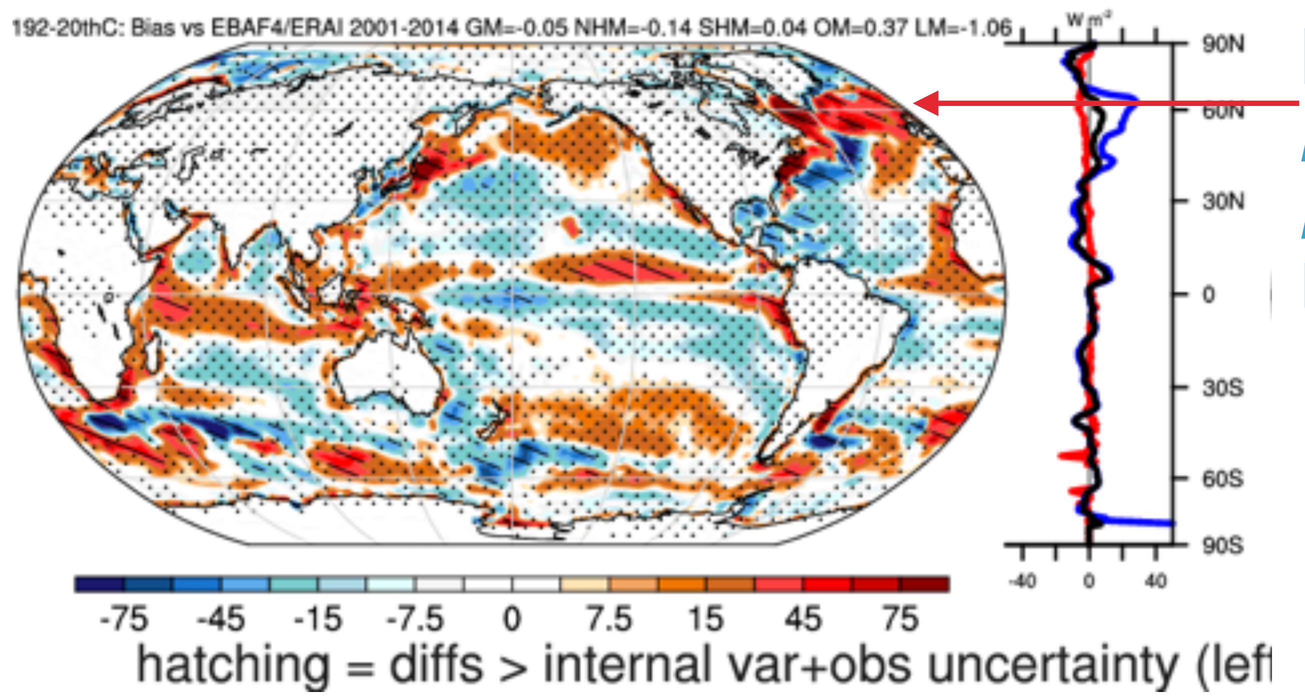


hatching = diffs > internal var+obs uncertainty (left column only); stippling = diffs > internal var only

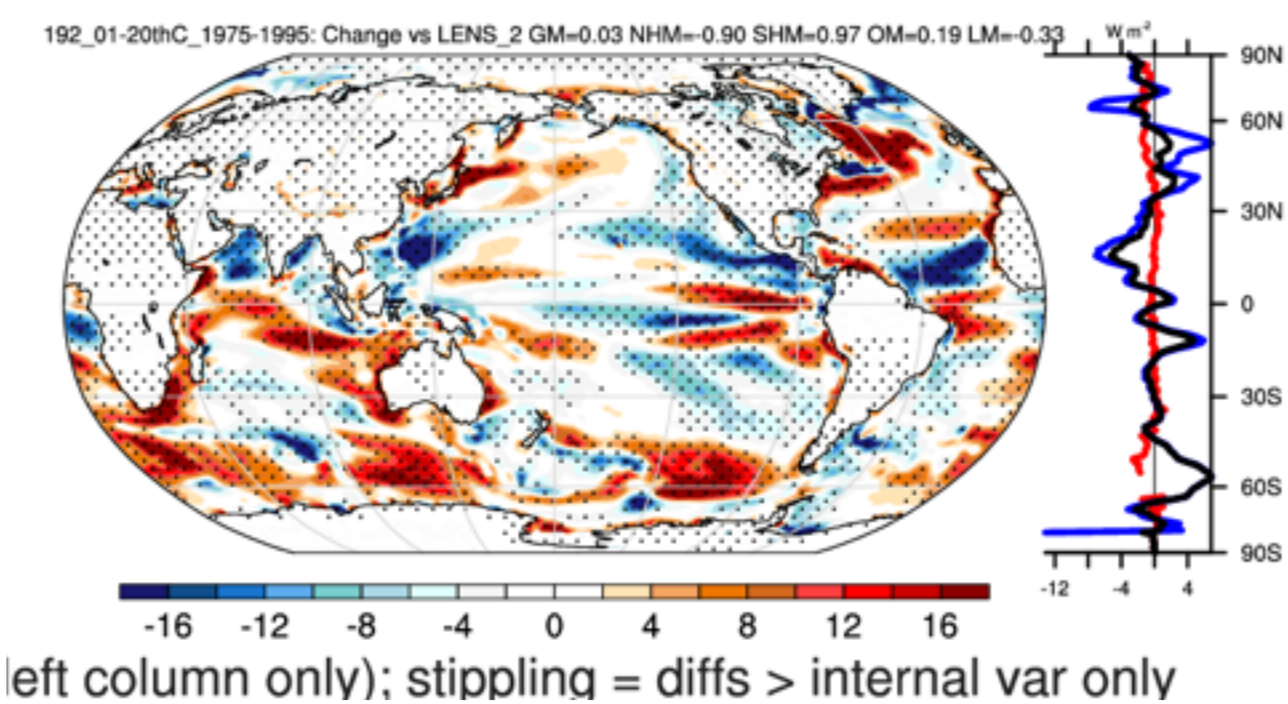
WHEN GHGS INTERRUPT THIS FEEDBACK MODEL WARMS RAPIDLY (1990S)

BIASED ARCTIC FEEDBACK: CONSEQUENCES FOR N. ATLANTIC

NET UPWARD SURFACE FLUX : 192 ANNUAL MEAN



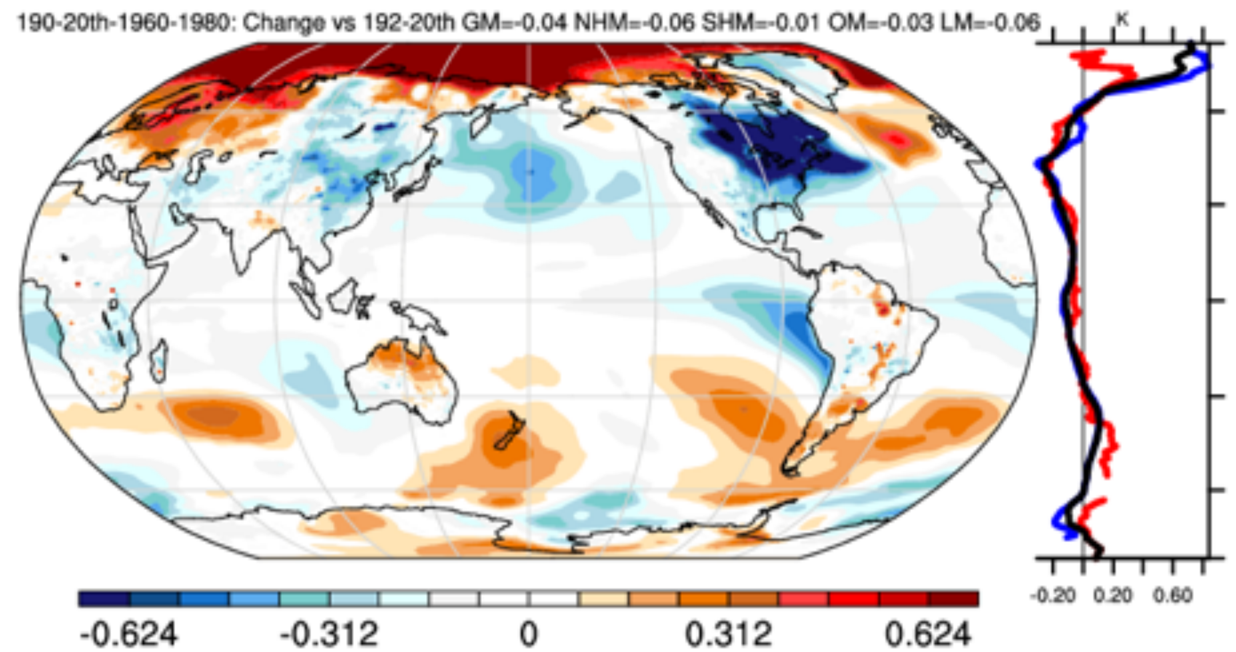
BIASED BY 45–100 W/M2 ACROSS NORTH ATLANTIC IN ANNUAL MEAN IN 192: UPSTREAM EFFECTS?

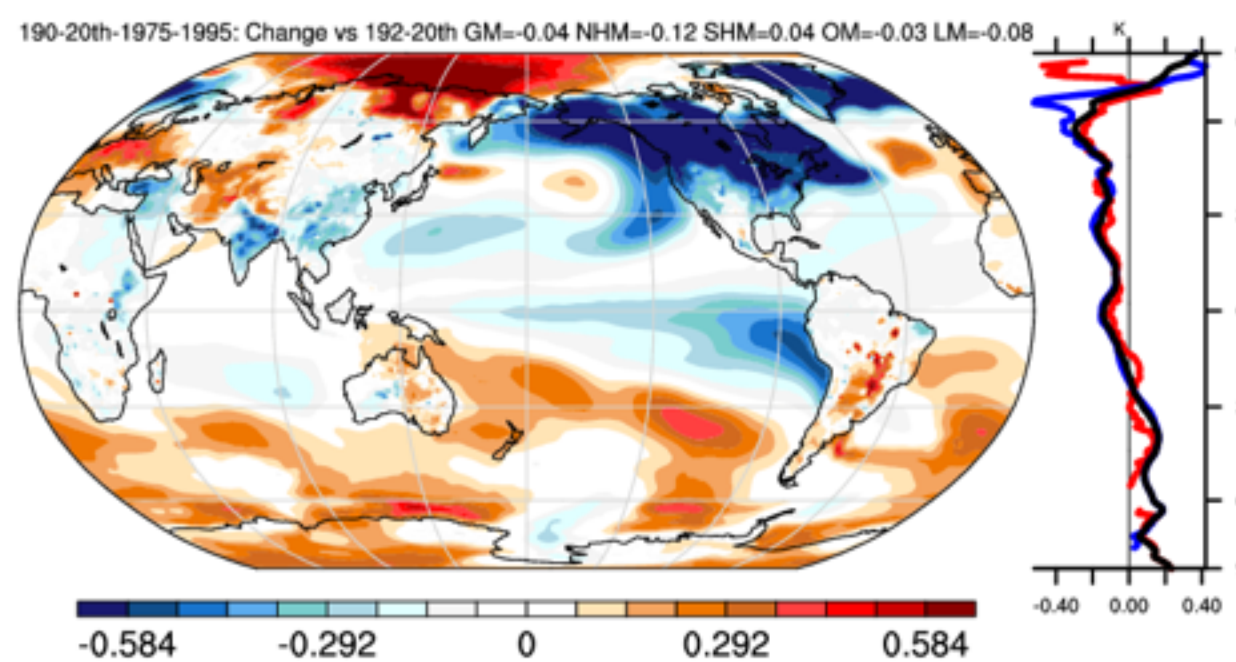


BIASES ARE WORSE THAN IN LENS BY >20 W/M2 IN LAB SEA/N ATL

BIASED ARCTIC FEEDBACK: TEMPERATURE PATTERNS

TEMPERATURE DIFFERENCES BETWEEN 190/192





TEMPERATURE DIFFERENCES LARGEST IN NORTH HEM / NORTH AMERICA

1960–1980

1975–1995

ON UTILITY OF CLOUD FORCING?

- Comparison between observed and modeling cloud forcing is inherently limited - observations are biased toward quiescent regions and are rarely cloud free

- But all sky fluxes also have limitations - they reflect multiple influences and the right flux can be simulated for the wrong reason.

- Clouds are the dominant uncertainty and driver of the present day energy budget and future changes. They deserve heightened scrutiny.

- So long as model error is large compared to the caveats in the observations, cloud forcing can serve as a useful constraint.

UTILITY OF TOA ENERGY BALANCING

- Tuning aspects of the atmospheric model is typically done before a run is begun so that net TOA flux (RESTOM) ~ 0 to minimize impacts on a transient (i.e. 20th C) simulation
- BUT - the ocean can have large drifts even with RESTOM=0 through compensation between the deep ocean and upper ocean, or regionally. Drifts in the upper ocean can have a large influence on transient (e.g. 20th C) runs. There is therefore a need to know how the upper ocean layers are drifting, more than whether RESTOM=0 or ΔOHC=0.